CANONICAL FORMS FOR LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEMS

By

PRAMOD PRABHAKAR KHARGONEKAR

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1981

Bound by DOBBS BROS. LIBRARY BINDING CO., INC., St. Augustine, Florida

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to all those who contributed towards making this work possible.

Professor R. E. KALMAN, the chairman of my supervisory committee, has been a constant source of encouragement during my entire graduate work. The main motivation for this work was provided by his research proposal and several invaluable discussions with him on linear-quadratic optimal control problems. I am also grateful to him for his early intuition of the relevance of the Riccati transformation group to the spectral theory of linear-quadratic optimal control problems. Without the financial support which he arranged for me for the last three years, this work would not exist today.

I would like to express my sincere appreciation to the other members of my supervisory committee: Professors T. E. BULLOCK, R. L. LONG, V. M. POPOV, and C. V. SHAFFER.

Discussions with Professor T. E. BULLOCK on the linear-quadratic optimal control problem and with Professor R. L. LONG on invariant theory and algebraic geometry helped a great deal in gaining an understanding of the problems considered in this dissertation.

A doctorate is more than a dissertation--it is a slow educational process. I gratefully acknowledge Professors E. EMRE, E. D. SONTAG, E. W. KAMEN, A. TANNENBAUM, G. BASILE, and others for their guidance and help.

I would also like to express my deep appreciation of the untiring help and unending patience of my fellow students and friends Tryphon, Bulent, Jaime, and Satish.

I am also grateful to Ms. Eleanor Onoda for doing an excellent typing job.

Of course, no research would be reported here were it not for the constant love and encouragement of my parents. To them, I am eternally grateful.

This research was supported in part by US Army Research Grant DAAG29-80-C0050 and US Air Force Grant AFOSR76-3034D through the Center for Mathematical System Theory, University of Florida, Gainesville, FL 32611, USA.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

ABSTRACT v

CHAPTER

I. INTRODUCTION 1

II. RICCATI TRANSFORMATION GROUP 5

1. The Linear-Quadratic Optimization Problem 5
2. Riccati Transformation Group 8
3. Invariants, Group Actions, and Canonical Forms . . . 14

III. STATE SPACE AND FEEDBACK GROUPS 17

4. Invariants for the State Space Group Action 18
5. Action of the Feedback Group 24

IV. ACTION OF THE RICCATI GROUP--I 31

6. Finite Time Problems--I 31
7. The Orthogonal Group 36
8. Finite Time Problems--I 39
9. Infinite Time Zero Terminal State Case 45

V. ACTION OF THE RICCATI GROUP--II 51

10. The Algebraic Riccati Equation Approach 51

VI. SYSTEM THEORETIC APPLICATIONS 63

11. Spectral Theory of Linear Quadratic Optimal Control. 63
12. Other Applications 80

REFERENCES 86

BIOGRAPHICAL SKETCH 91

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

CANONICAL FORMS FOR LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEMS

By

PRAMOD PRABHAKAR KHARGONEKAR

August, 1981

Chairman: Dr. R. E. Kalman
Cochairman: Dr. T. E. Bullock
Major Department: Electrical Engineering

This work is concerned with the Riccati group of transformations acting on spaces of various types of linear optimal control problems with quadratic cost criteria. Canonical forms and complete systems of invariants are studied for this group action and several results are obtained. These canonical forms and the Riccati transformation group are shown to have several important system theoretic applications. In particular, spectral theory of linear-quadratic optimal control problems, Riccati differential equation, positive real transfer functions are investigated via the Riccati transformation group. A complete solution to the problem of existence of solution to the so called "infinite time free terminal state" problems is given in terms of non-negativity of a symmetric matrix. This solution is also related to the "minimal energy" control problem.

CHAPTER I. INTRODUCTION

Linear-quadratic optimal control problem is one of the most fundamental problems of modern system theory. Optimal control and filtering problems led to the concepts of reachability and observability--the key concepts in system theory today. The pioneering work on linear-quadratic optimal control theory was done by KALMAN [1960]. Some other papers which made substantial contributions to the theory of linear-quadratic optimal control include POPOV [1964], WILLEMS [1971], and JONCKHEERE and SILVERMAN [1978]. There are several text books on the subject of which we mention BROCKETT [1970], ANDERSON and MOORE [1971], and WILLEMS [1981]. Despite the voluminous research on this problem, some aspects of the theory have not been fully investigated. This dissertation explores one such direction.

Very often, mathematical objects admit a group of transformations which corresponds to some notion of equivalence among these objects. Action of $GL(n)$ (the set of $n \times n$ nonsingular matrices) on the space of $n \times n$ matrices by conjugation is a very good example of this. The underlying notion of equivalence in this case arises from change of basis. The main mathematical problem in such a situation is to find quantities which remain invariant under the transformation. In the example above, the invariant factors and eigenvalues are some of the invariant quantities. The basic idea is that these invariant quantities reveal the fundamental features of the mathematical objects under consideration.

Mathematical system theory provides examples of this type of problems. For example, $GL(n)$ acts on the set of reachable systems in a natural way. This action corresponds to a change of basis in the state space. Again a detailed investigation of this problem from an invariant theoretic point of view has revealed the basic structure of reachable systems. See POPOV [1972], KALMAN [1974], TANNENBAUM [1980] for an account of this and some associated problems. (In Section 4, we briefly discuss some of the results.)

The main problem of this dissertation arises from analogous considerations. A group of transformations which acts naturally on linear-quadratic optimal control problems was first formulated by KALMAN [1975] in a

research proposal. This group of transformations soon became known as the Riccati group. Roughly speaking, the Riccati group is built out of four components: state space transformations, input space transformations, linear state variable feedback, and change of variables in the Riccati differential equation. The main problem of this dissertation is to find invariants and canonical forms for linear-quadratic optimal control problems under the action of the Riccati group. Our results of Chapters IV and V show that we are substantially successful in resolving this problem. The results are very satisfactory for the case of equal Kronecker indices. We also treat the case of strictly unequal Kronecker indices explicitly. The results in the general case do not seem to be easy to formulate. However, in any concrete situation the techniques developed in this dissertation can be used to obtain invariants and canonical forms.

We will now briefly describe the contents of the following chapters.

Chapter II is devoted to a precise formulation of the Riccati transformation group. We analyze each "component" of the Riccati group separately, and combine all the four components into a single transformation group. The basic ideas of this chapter are essentially contained in KALMAN [1975].

The purpose of Chapter III is to review some well known material on the action of the state space group and the feedback group on the space of reachable pairs. We define a new canonical form for the action of the feedback group called "dual Kronecker canonical form". This canonical form is based on the duality of the associated Young's diagrams. There are two main contributions in Chapter III. We show that the set of reachable systems with a given set of unordered Kronecker indices is an affine variety; a geometric quotient for the action of the general linear group on this space exists, the geometric quotient turns out to be a Euclidean space, and the quotient map is given by the well known invariants due to POPOV [1972]. This result should be viewed as an algebro-geometric reformulation of the results of POPOV [1972]. In Section 5, we formulate the stabilizer in the feedback group of a reachable pair in dual Kronecker canonical form. Previously, WANG and

DAVISON [1976] had formulated the stabilizer in the feedback group of a reachable pair in the Kronecker canonical form.

Chapter IV contains the first results on the invariants and canonical forms for the action of the Riccati group. The results are very satisfactory for the case of equal Kronecker indices. We also treat the case of strictly unequal Kronecker indices explicitly. Some of the results of this chapter turn out to be useful in Chapter VI on system theoretic applications.

In Chapter V, we develop another approach to obtaining invariants and canonical forms. Here we assume that the associated algebraic Riccati equation has a real symmetric solution. Again, the results are very satisfactory for the case of equal Kronecker indices. The difficulties in the general case are analogous to the difficulties in Chapter IV.

Chapter VI contains some system theoretic applications of the results of Chapters IV and V. The main application is to the spectral theory of linear-quadratic optimal control problems. This theory was developed by JONCKHEERE and SILVERMAN [1978] for the discrete-time case. They analyze the so called "infinite time free terminal state linear-quadratic optimal control problem" from a Hilbert space operators point of view. The problem of boundedness of infima is reduced to the nonnegativity of a certain self-adjoint Hilbert space operator. However, it turns out that a direct calculation of the spectrum of this operator is not tractable.

We consider the continuous-time analog of the problem considered by JONCKHEERE and SILVERMAN [1978] as formulated originally by WILLEMS [1971]. We show that the existence of solution to the optimization problem remains invariant under the action of the Riccati group. However, the Hilbert space operator defined by JONCKHEERE and SILVERMAN [1978] does not remain invariant under the Riccati group. We first compute a Hilbert space operator which is the continuous-time analog of the operator defined by JONCKHEERE and SILVERMAN [1978]. Use of the Riccati group leads to a major simplification in the derivation of this operator. We then show using the Riccati group that this operator can be brought to a very simple form. In this form, the nonnegativity of this operator becomes equivalent to the nonnegativity of a certain symmetric matrix. Thus, the problem

of boundedness of infima can be reduced to checking nonnegativity of a symmetric matrix. We rederive a generalized version of this result using certain simple results from the theory of "minimal energy control" developed by KALMAN, HO, and NARENDRA [1963]. This gives a very simple system theoretic interpretation of the spectral theoretic results.

We also examine Popov's frequency function, positive real matrices, and fast algorithms for the Riccati differential equation due to KAILATH [1973] and CASTI [1974] in the light of the Riccati transformation group. Under certain mild conditions, we completely parametrize scalar positive real functions. Further work on the fast algorithms from the Riccati group point of view seems promising. Generalization of our results on scalar positive real functions to the matrix case is also of considerable interest.

We close this introduction with a few words on notation and standing assumptions. Throughout this dissertation, we will work with the field of real numbers. $\underline{\underline{R}}^m$ denotes the m-dimensional real vector space consisting of m-length column vectors with real entries. $\underline{\underline{R}}^{m\times n}$ denotes the set of $m \times n$ matrices over the field of real numbers. $L^2[a, b]$ denotes the Hilbert space of square integrable functions on the interval [a, b]. A polynomial χ is called a Hurwitz polynomial iff the roots of χ are in the open left half plane. All undefined terminology from invariant theory and algebraic geometry is standard for which the reader is referred to FOGARTY [1965], MUMFORD [1965], TANNENBAUM [1980].

CHAPTER II. RICCATI TRANSFORMATION GROUP

In this chapter, we shall introduce the Riccati transformation group, which arises naturally as a group of transformations of linear-quadratic optimization problems. This group was originally formulated by KALMAN [1975] in a research proposal. Roughly speaking, the Riccati transformation group consists of state space transformations, input space transformations, linear state variable feedback, and change of variables in the Riccati equation. We shall discuss each of these transformations separately and then combine all these into a single group action.

1. The Linear Quadratic Optimization Problem.

The "dynamic optimization problem with integral quadratic criteria" for a continuous-time, finite dimensional, linear system is defined as follows: minimize the functional

$$(1.1)\quad J(u(\cdot), c, t_o) = \int_{t_o}^{o} [x'(\tau)Qx(\tau) + 2x'(\tau)Su(\tau) + u'(\tau)Ru(\tau)]d\tau + x'(0)P_o x(0),$$

over all continuous $\underline{\underline{R}}^m$ valued input functions $u(\tau)$, $0 \geq \tau \geq t_o$, with respect to the solutions to the n-th order continuous-time linear system

$$(1.2)\quad \frac{dx(\tau)}{d\tau} = Fx(\tau) + Gu(\tau), \quad x(t_o) = c.$$

Here $x(\tau)$ is in $\underline{\underline{R}}^n$, F, G, Q, R, S, P_o are respectively $n \times n$, $n \times m$, $n \times n$, $m \times m$, $n \times m$, $n \times n$ real matrices, Q, R, P_o are assumed to be symmetric matrices. R is always assumed to be positive definite. We also assume that (F, G) is reachable. It should be noted that the system as well as the quadratic cost criteria are time-invariant. It is easy to derive results analogous to our results which apply to the discrete-time linear quadratic optimization problem: minimize the functional

$$(1.3)\quad J(u(\cdot),\ c,\ N) = \sum_{k=-N}^{1}[x'(k)Qx(k) + 2x'(k)Su(k) + u'(k)Ru(k)] + \\ + x'(0)P_o x(0)$$

over all $\underline{\underline{R}}^m$ valued input sequences $u(k)$, $0 \geq k \geq -N$, with respect to the n-th order discrete-time linear system.

$$(1.4)\quad x(k+1) = Fx(k) + Gu(k),\quad x(-N) = c.$$

Here F, G, Q, R, S, P_o, $x(k)$ are as before. The only difference is that R is not assumed to be positive definite.

The linear quadratic-optimization problems defined above have been studied extensively over the last two decades. The literature surrounding these problems has grown enormously. The reader is referred to the seminal paper of KALMAN [1960]. See also KALMAN [1964], POPOV [1964], WILLEMS [1971], JONCKHEERE and SILVERMAN [1978], and the references cited there.

It is well known that the solution to the above continuous-time problem is based on the (matrix) Riccati differential equation. The optimal control input $u_*(\cdot)$ is given by

$$u_*(\tau) = L(\tau)x_*(\tau),\quad t_o \leq \tau \leq 0,\quad \frac{dx_*(\tau)}{d\tau} = Fx_*(\tau) + Gu_*(\tau),\quad x_*(t_o) = c,$$

where

$$(1.5)\quad L(\tau) = -R^{-1}(G'P(\tau) + S'),$$

and $P(\cdot)$ is the solution of the Riccati differential equation

$$(1.6)\quad \frac{-dP(\tau)}{d\tau} = F'P(\tau) + P(\tau)F + Q - (P(\tau)G + S)R^{-1}(P(\tau)G + S)',\quad P(0) = P_o$$

Here we implicitly assume that the solution to (1.6) exists.

The problem of existence of solution to (1.6) is, in general, a difficult problem. Under certain positivity assumptions on the cost criteria J, the existence of solutions to (1.6) can be guaranteed. See

Chapter VI for a discussion of the general case. A similar feedback solution based on the Riccati difference equation can be given for the discrete-time problem.

We thus see that the optimization problem as well as its solution can be fully described in terms of the "problem data" F, G, Q, R, S, P_o. (Since the initial state c plays no special role in various features of the optimization problem, we do not consider it in the "problem data".) Thus, every linear quadratic optimization problem can be identified with a six tuple (F, G, Q, R, S, P_o). Conversely, if (F, G, Q, R, S, P_o) is any six tuple of matrices of appropriate sizes, (F, G) is reachable, Q, R, P_o are symmetric, and R is positive definite then we can define a linear quadratic optimization problem associated with (F, G, Q, R, S, P_o). In view of these elementary observations, we define the (finite time) optimization data space

$$(1.7)\quad D := \{(F, G, Q, R, S, P_o) \text{ in } \underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m} \times \underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{m\times m} \times \underline{\underline{R}}^{n\times m} \times \underline{\underline{R}}^{n\times n}: (F, G) \text{ reachable, } Q, R, P_o \text{ symmetric, and } R \text{ positive definite}\}.$$

Similar considerations also apply to the discrete-time linear quadratic optimization problems. In particular, we have discrete-time optimization data space $D_d := \{(F, G, Q, R, S, P_o)$ in $\underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m} \times \underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{m\times m} \times \underline{\underline{R}}^{n\times m} \times \underline{\underline{R}}^{n\times n}$: (F, G) reachable, Q, R, P_o symmetric$\}$.

An important problem closely related to the linear-quadratic optimization problem is the so-called <u>infinite time zero terminal state linear-quadratic optimal control problem</u> defined as follows.

Minimize

$$(1.8)\quad J(u(\cdot), c) = \int_o^\infty [x'(\tau)Qx(\tau) + 2x'(\tau)Su(\tau) + u'(\tau)Ru(\tau)]d\tau$$

subject to the linear system

$$\frac{dx(\tau)}{d\tau} = Fx(\tau) + Gu(\tau), \quad x(0) = c, \quad \lim_{\tau\to\infty} x(\tau) = 0,$$

over all possible $\underline{\underline{R}}^m$-valued, continuous, square integrable input functions $u(\cdot)$. Necessary and sufficient conditions for the existence of solution to this problem are well known. In particular, a solution to the above infinite time zero terminal state problem exists if and only if the algebraic Riccati equation

$$(1.9) \quad F'P + PF + Q - (PG + S)R^{-1}(PG + S)' = 0$$

admits a real symmetric solution. In case (1.9) has a real symmetric solution, it has a unique real symmetric solution P such that the eigenvalues of $(F - GR^{-1}(G'P + S'))$ are in the (closed) left half plane; i. e.,

$$\mathrm{Re}(\lambda(F - GR^{-1}(G'P + S'))) \leq 0.$$

WILLEMS [1971] is an excellent reference for a detailed analysis of the algebraic Riccati equation.

It is clear from the above discussion that the infinite time zero terminal state problems are completely specified by the quintuple of matrices (F, G, Q, R, S). In view of this we define the optimization data space for the infinite time zero terminal state problems

$$DS := \{(F, G, Q, R, S) \text{ in } \underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m} \times \underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{m\times m} \times \underline{\underline{R}}^{n\times m}:$$

$$(F, G) \text{ reachable, } Q, R \text{ symmetric, } R \text{ positive definite}\}.$$

Analogous considerations apply to the discrete-time case.

2. Riccati Transformation Group.

The purpose of this section is to introduce some transformations of the above dynamic optimization problems which arise naturally from the problem statement and the associated feedback solution based on the Riccati differential equation.

A. Change of basis in the state space. Let T be any $n \times n$ real nonsingular matrix; i. e., T is in $GL(n)$. If we introduce the new state variables

$$z(t) := T^{-1}x(t),$$

then the linear quadratic optimization problem (1.1) transforms to minimize

$$J(u(\cdot), T^{-1}c, t_o) = \int_{t_o}^{o}[z'(\tau)T'QTz(\tau) + 2z'(\tau)T'Su(\tau) + u'(\tau)Ru(\tau)]d\tau + z'(0)T'P_oTz(0),$$

subject to the linear system

$$\frac{dz(\tau)}{d\tau} = T^{-1}FTz(\tau) + T^{-1}Gu(\tau), \quad z(t_o) = T^{-1}c.$$

Thus, the problem data (F, G, Q, R, S, P_o) transform to $(T^{-1}FT, T^{-1}G, T'QT, R, T'S, T'P_oT)$. Furthermore, $P(\tau)$ is the solution to the Riccati differential equation associated with (F, G, Q, R, S, P_o) if and only if $T'P(\tau)T$ is the solution to the Riccati differential equation associated with $(T^{-1}FT, T^{-1}G, T'QT, R, T'S, T'P_oT)$. It is easy to check that the optimal control inputs for the original problem and the transformed problem are the same. Consequently, state space basis changes do not alter the basic features of the linear quadratic optimization problems. Summarizing, $GL(n)$ acts on the data space D by

$$(2.1) \quad \varphi_T: D \times GL(n) \to D: ((F, G, Q, R, S, P_o), T) \mapsto \mapsto (T^{-1}FT, T^{-1}G, T'QT, R, T'S, T'P_oT).$$

Finally, we note that the action of $GL(n)$ on the space of reachable pairs (F, G) has been considered from several points of view in the literature. See, for example, POPOV [1972], KALMAN [1974], TANNENBAUM [1980] and the reference cited there. (We summarize some of these results in Section 4.) The main difference, of course, is that there is an associated optimization problem.

B. Change of basis in the input space. This transformation is similar to the state space transformation discussed above. Let V be an $m \times m$ nonsingular matrix. We introduce new input variables $v(\cdot)$ in $\underline{\underline{R}}^m$ by

$$v(\tau) = V^{-1}u(\tau).$$

Proceeding as in A, this leads to the transformation law

$$(2.2) \quad \varphi_V: D \times GL(m) \to D: ((F, G, Q, R, S, P_o), V) \mapsto$$

$$\mapsto (F, GV, V'RV, SV, P_o).$$

C. Linear state variable feedback. This is perhaps the most nontrivial transformation of the original problem. Suppose we choose

$$\hat{u}(\tau) := u(\tau) - Lx(\tau),$$

where L is an $m \times n$ matrix and $u(\cdot)$ is any continuous $\underline{\underline{R}}^m$-valued function on $[t_o, 0]$. The linear system then transforms to

$$\frac{dx(\tau)}{d\tau} = (F + GL)x(\tau) + G\hat{u}(\tau), \quad x(t_o) = c.$$

We also have

$$x'(\tau)Qx(\tau) + 2x'(\tau)Su(\tau) + u'(\tau)Ru(\tau) = x'(\tau)[Q + SL + L'S' + L'RL]x(\tau) +$$

$$+ 2x'(\tau)(S + L'R)x(\tau) + v'(\tau)Rv(\tau).$$

Thus the original problem data $d := (F, G, Q, R, S, P_o)$ transform to $d_L := (F + GL, G, Q + L'RL + SL + L'S', R, S + L'R, P_o)$. It is a routine verification that $P(\tau)$ is a solution to the Riccati differential equation associated with d if and only if $P(\tau)$ is the solution to the Riccati differential equation associated with d_L, (whenever they exist). The optimal feedback matrices $L(d)$ and $L(d_L)$ are related by

$$(2.3) \quad L(d_L)(\tau) = L(d)(\tau) - L.$$

Summarizing, linear state variable feedback induces the transformation

$$(2.4)\quad \Phi_L: D \times \underline{\underline{R}}^{m\times n} \to D: ((F, G, Q, R, S, P_o), L) \mapsto$$

$$\mapsto (F + GL, G, Q + SL + L'S' + L'RL, R, S + L'R, P_o).$$

It has been recognized previously in literature that the feedback transformation does not alter most of the basic features of the dynamic optimization problem. See, for example, SILVERMAN [1976], MOLINARI [1975], JONCKHEERE and SILVERMAN [1978], etc. Feedback transformation is used to stabilize the system by JONCKHEERE and SILVERMAN [1978]. As we have shown above, the form of the linear quadratic optimization problem does not alter under feedback, solution to the Riccati differential equation does not alter, and the optimal feedback matrices are related via a simple translation by L as in (2.3).

D. Change of variables in the Riccati differential equation. This transformation arises from the fact that the solution to the linear quadratic optimization problem is given in terms of the Riccati differential equation

$$(1.6)\quad \frac{-dP(\tau)}{d\tau} = F'P(\tau) + P(\tau)F + Q - (P(\tau)G + S)R^{-1}(P(\tau)G + S)',$$

$$P(0) = P_o.$$

Suppose we perform the change of variables

$$\hat{P}(\tau) := P(\tau) - M$$

where M is an $n \times n$ constant symmetric matrix. For $\hat{P}(\tau)$, we obtain the differential equation

$$(2.5)\quad \frac{-d\hat{P}(\tau)}{d\tau} = F'\hat{P}(\tau) + \hat{P}(\tau)F + (Q + F'M + MF) - (\hat{P}(\tau)G + S + MG)$$

$$R^{-1}(\hat{P}(\tau)G + S + MG)', \quad \hat{P}(0) = P_o - M.$$

As is clear from this transformed equation, (2.5) is the Riccati differential equation associated with the data $(F, G, Q + F'M + MF, R, S + MG, P_o - MG)$. Furthermore, the cost associated with this transformed problem is

$$J(u(\cdot), c, t_o) := \int_{t_o}^{o} [x'(\tau)(Q + F'M + MF)x(\tau) + 2x'(\tau)(S + MG)u(\tau) + u'(\tau)Ru(\tau)]d\tau + x'(0)(P_o - M)x(0).$$

Recalling that $x(\tau)$ is governed by

$$\frac{dx(\tau)}{d\tau} = Fx(\tau) + Gu(\tau),$$

we have

$$\begin{aligned} J_M(u(\cdot), c, t_o) &= \int_{t_o}^{o} [x'(\tau)Qx(\tau) + 2x'(\tau)Su(\tau) + u'(\tau)Ru(\tau) + \\ &\quad + \frac{d}{d\tau}(x'(\tau)Mx(\tau))]d\tau + x'(0)P_o x(0) - x'(0)Mx(0) \\ &= J(u(\cdot), c, t_o) - c'Mc. \end{aligned}$$

Thus, for every fixed initial state c, an infima for J_M exists if and only if an infima for J exists. Further, an input function $u(\cdot)$ minimizes J_M if and only if $u(\cdot)$ minimizes J, for each fixed initial state c. Finally, $\hat{P}(\tau)$ is a solution to (2.5) if and only if $\hat{P}(\tau) + M$ is a solution to (1.6), and the associated optimal feedback gain matrices are the same. Thus, this transformation also preserves most of the basic features of the problem. To conclude, let $Sym(n)$ be the additive group of $n \times n$ symmetric matrices. We then have the transformation

$$\begin{aligned} (2.6) \quad \varphi_M: D \times Sym(n) \to D: ((F, G, Q, R, S, P_o), M) &\mapsto \\ &\mapsto (F, G, Q + F'M + MF, R, S + MG, P_o - M). \end{aligned}$$

This transformation occurs in several different contexts in system theory literature. In the study of stability and optimality this transformation arises in POPOV [1973]. In the study of stochastic realization problem, this transformation has been considered by CANDY [1975]. In the closely related problem of positive real matrices, this transformation occurs in ANDERSON [1967]. BERNHARD and COHEN [1973] consider a similar transformation for the inverse problem of optimal control.

We will now combine the four different types of transformation discussed above. This will constitute the action of the Riccati group on the data space D or DS. Consider the set

$$\Gamma := \{(T, V, L, M): T \text{ in } GL(n),\ V \text{ in } GL(m),\ L \text{ in } \underline{\underline{R}}^{m\times n},\ M \text{ in } Sym(n)\}. \tag{2.7}$$

We define a binary operation $\circ$ between any two elements of Γ given by

$$(T_1, V_1, L_1, M_1) \circ (T_2, V_2, L_2, M_2) :=$$

$$:= (T_1T_2, V_1V_2, V_1L_2 + L_1T_2, M_1 + (T_1')^{-1}M_2T').$$

It is easy to check that Γ with the binary operation $\circ$ is a group. We just note that $e := (I, I, 0, 0)$ serves as the identity and the inverse of (T, V, L, M) is $(T^{-1}, V^{-1}, -V^{-1}LT, -T'MT)$. We now define the map

$$\varphi: D \times \Gamma \to D: ((F, G, Q, R, S, P_o), (T, V, L, M)) \mapsto \tag{2.8}$$

$$\mapsto (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o),$$

where

$$\hat{F} := T^{-1}FT + T^{-1}GL,$$

$$\hat{G} := T^{-1}GV,$$

$$\hat{Q} := T'(Q + F'M + MF)T + T'(S + MG)L + L'(S + MG)'T + L'RL,$$

$$\hat{R} := V'RV,$$

$$\hat{S} := (T'(S + MG) + L'R)V,$$

$$\hat{P}_o := T'(P_o + M)T.$$

In an entirely analogous manner, we can define the action of the Riccati group Γ on the space DS corresponding to the infinite-time zero terminal state problems as

$$\psi\colon DS \times \Gamma \to DS\colon ((F, G, Q, R, S), (T, V, L, M)) \mapsto (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S})$$

where $\hat{F}$, $\hat{G}$, $\hat{Q}$, $\hat{R}$, $\hat{S}$ are as above.

It is easy to check that the function φ satisfies the following rules: for any γ_1, γ_2 in Γ and d in D,

$$(2.9)\quad \varphi(d, \gamma_1 \circ \gamma_2) = \varphi(\varphi(d, \gamma_1), \gamma_2)$$

$$(2.10)\quad \varphi(d, e) = d.$$

In fact the binary operation $\circ$ in Γ was defined precisely so that (2.9) and (2.10) hold.

3. Invariants, Group Actions, and Canonical Forms.

Let W be an abstract denoted multiplicatively and Z be a set. We say that W acts on Z (on the right) iff there exists a map $\mu\colon Z \times W \to Z$ such that for any w_1, w_2 in W and z in Z, we have

$$\mu(z, w_1w_2) = \mu(\mu(z, w_1), w_2), \quad \mu(z, e) = z.$$

(Here e is the identity of W.) The triple (W, Z, μ) is called an algebraic transformation space and μ is called the (right) action of W on Z. For any z in Z, $z \cdot W := \{\mu(z, w)\colon w \text{ in } W\}$ is called the orbit of z. Also, for any z in Z, we define the stabilizer of z to be $W_z := \{w \text{ in } W\colon \mu(z, w) = z\}$; i. e., the stabilizer of z consists of those group elements which leave z fixed. It is easy to check that the stabilizer W_z is a subgroup of W.

Now let Y be any other set and $f: Z \to Y$ be a function. Then f is said to be invariant (under the action of W) iff for any z in Z and w in W, $f(\mu(z, w)) = f(z)$. An invariant $f: Z \to Y$ is called a complete invariant iff for any z_1, z_2 in Z, $f(z_1) = f(z_2)$ implies $z_1 = \mu(z_2, w)$, for some w in W. The set Y is said to parametrize the orbit space iff f is a complete invariant and f is surjective. In this set theoretic setup, there is a trivial solution to the problem of finding a set Y that parametrizes the orbit space. One can choose Y to be the set of orbits of z in Z. However, such a solution, in general, is not very useful. It is more desirable to have S with some mathematical structure.

A group W is called an algebraic group iff (i) W is an algebraic variety, (ii) the two maps $\alpha: W \times W \to W: (w_1, w_2) \mapsto w_1 w_2$ and $\beta: W \to W: w \mapsto w^{-1}$ are morphisms. Now suppose W is algebraic group and Z is an algebraic variety such that the map

$$\mu: Z \times W \to W: (z, w) \mapsto \mu(z, w)$$

is a morphism. We then say that W acts on Z morphically. In such a situation, we may require the parametrizing set to be an algebraic variety. This problem is one of the central problems of the modern geometric invariant theory pioneered by MUMFORD [1965]. TANNENBAUM [1980] is an excellent reference for an expository account of geometric invariant theoretic aspects of linear system theory.

The action φ of Γ is a right action on D follows directly from (2.9) and (2.10). However, there are certain difficulties in directly applying the machinery of geometric invariant theory to the action of Riccati group on D. The main difficulty lies in the fact that D does not seem to have any algebraic geometric structure. The principal reason for this seems to be the condition $R > 0$. The positivity of a symmetric matrix is well known to be equivalent to certain polynomial inequalities (in the entries of R). Thus, it does not seem to be possible to apply the results of geometric invariant theory directly. In any case, it is easy to prove that Γ is a linear algebraic group.

However, whether Γ is linearly reductive seems to be a difficult question as well.

The next three chapters are concerned with answering questions on invariants and canonical forms for the action of Γ on D. We do use some aspects of modern invariant theory, but only after certain preliminary reductions of the original problem to somewhat more manageable situations.

CHAPTER III. STATE SPACE AND FEEDBACK GROUPS

The purpose of this chapter is to review some relevant results about the action of the state space and feedback groups on the space of reachable systems. The results described here will be the main preliminary material for obtaining invariants and canonical forms for the action of the Riccati group. This chapter is divided into two sections: first we discuss invariants and canonical forms for the action of the state space group, and next we discuss the action of the feedback group. Most of these results have been known for the past several years. We present a new result on the action of the state space group which may be viewed as a geometric counterpart of some previously known results on invariants due to POPOV [1972].

Let us define the space of reachable systems

$$\Sigma := \{(F, G): F \text{ in } \underline{\underline{R}}^{n\times n}, \ G \text{ in } \underline{\underline{R}}^{n\times m}, \ (F, G) \text{ reachable}\};$$

i. e., Σ denotes the set of all reachable systems with n states and m inputs. As in the preceding chapter, the group of state space basis changes $GL(n)$ acts on Σ in the following way:

$$\chi: \Sigma \times GL(n) \to \Sigma: ((F, G), T) \mapsto (T^{-1}FT, T^{-1}G).$$

It is easy to verify that χ is a right action of $GL(n)$ on Σ. The feedback group Γ_f is defined as

$$\Gamma_f := \{(T, V, L): T \text{ in } GL(n), \ V \text{ in } GL(m), \ L \text{ in } \underline{\underline{R}}^{m\times n}\}.$$

Γ_f is a group under the multiplication $\circ$ given by

$$(T_1, V_1, L_1)\circ(T_2, V_2, L_2) = (T_1T_2, V_1V_2, L_1T_2 + V_1L_2)$$

Γ_f acts on Σ according to

$$\theta: \Sigma \times \Gamma_f \to \Sigma: ((F, G,), (T, V, L)) \mapsto (T^{-1}FT + T^{-1}GL, T^{-1}GV).$$

θ is called the <u>feedback action</u>. It is easy to verify that θ is a

right action of Γ_f on Σ. The groups $GL(n)$ and Γ_f can be naturally identified with subgroups of the Riccati group; in fact, this is the main reason that leads us to study actions of these groups on the space of reachable systems.

4. Invariants for the State Space Group Action.

The action χ of $GL(n)$ on Σ has been studied extensively from several points of view. A complete set of invariants, canonical forms, and a geometric quotient have been discovered. We shall describe some of these results very briefly. For detailed proofs and discussions, the interested reader is referred to POPOV [1972], KALMAN [1972, 1974], HAZEWINKEL and KALMAN [1976], HAZEWINKEL [1977], TANNENBAUM [1980], BYRNES and GAUGER [1977], and others.

Define the reachability matrix map

$$(4.1) \qquad \rho\colon \Sigma \to \underline{\underline{R}}^{n\times nm}\colon (F, G) \mapsto (G, FG, \ldots, F^{n-1}G) =: \rho(F, G)$$

Number the columns of $\rho(F, G)$ lexicographically (from left to right) as $01, \ldots, 0m, 11, \ldots, 1m, \ldots, (n-1)1, \ldots, (n-1)m$. A nice selection is defined to be a subset ν of these indices of size n such that if ij belongs to ν then ℓj belongs to ν for $\ell < i$. For a nice selection ν define

$$\rho_\nu\colon \Sigma \to \underline{\underline{R}}^{n\times n}\colon (F, G) \mapsto \rho_\nu(F, G),$$

where $\rho_\nu(F, G)$ denotes the $n \times n$ submatrix of $\rho(F, G)$ consisting of the columns indexed by the nice selection ν. One of the key results of KALMAN [1974] is that (F, G) is reachable if and only if $\rho_\nu(F, G)$ is nonsingular for some nice selection ν.

(4.2) REMARK. For each nice selection ν, define open affine subsets U_ν of $\underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m}$ by

$$U_\nu := \{(F, G) \text{ in } \underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m}\colon \det \rho_\nu(F, G) \neq 0\}, \quad (U_\nu \subseteq \Sigma).$$

It is easy to see that $GL(n)$ acts on Σ in such a way that $\chi(U_\nu) \subseteq U_\nu$; i. e., U_ν is stable under the action of $GL(n)$. Further, $\{U_\nu: \nu \text{ a nice selection}\}$ constitutes an affine open cover of Σ. TANNENBAUM [1980] shows that $GL(n)$ acts on Σ, and hence on each U_ν, with closed orbits. Using the geometric invariant theory of MUMFORD [1965], TANNENBAUM then proves that a geometric quotient V_ν for the action of $GL(n)$ on U_ν exists for each nice selection ν. Finally, using the techniques of modern algebraic geometry, it is shown that V_ν "patch up" in such a way that a geometric quotient for the action of $GL(n)$ on Σ exists. This geometric quotient turns out to be a smooth, irreducible, quasiprojective variety of dimension nm. However, we will not work directly with this geometric quotient.

Following POPOV [1972], we will now describe how a special nice selection, called the Kronecker nice selection, can be naturally associated with every reachable system (F, G) in Σ. Let (F, G) be a reachable pair. Consider the ordered set of vectors

$$\Lambda := \{g_1, g_2, \ldots, g_m, Fg_1, \ldots, Fg_m, \ldots, F^{n-1}g_1, \ldots, F^{n-1}g_m\},$$

where g_i is the i-th column of the matrix G. In other words, Λ consists of columns of $\rho(F, G)$ listed from the left to the right. A vector $F^i g_j$ is said to be an antecedent of another vector $F^p g_q$ iff $im + j < pm + q$; i. e., $F^i g_j$ occurs to the left of $F^p g_q$ in the list Λ. We can now define some well known invariants of the pair (F, G).

(4.3) DEFINITIONS. Let (F, G) be in Σ and j be an integer, $1 \leq j \leq m$. Then the j-th unordered Kronecker index κ_j of (F, G) is the smallest integer κ such that $F^\kappa g_j$ is linearly dependent upon its antecedents. Integers $\kappa_1, \kappa_2, \ldots, \kappa_m$ are called the unordered Kronecker indices of the pair (F, G). The integers $\kappa_1^*, \kappa_2^*, \ldots, \kappa_m^*$ obtained by reordering $\kappa_1, \kappa_2, \ldots, \kappa_m$ in a nonincreasing order are called the ordered Kronecker indices of (F, G). We note that the condition rank $G = m$ is equivalent to the

condition $\kappa_i \geq 1$, $i = 1, 2, \ldots, m$. We will often make this assumption for the sake of notational and computational simplicity. Let us define $\hat{\Sigma} := \{(F, G) \text{ in } \Sigma: \text{rank } G = m\}$.

Let (F, G) be in Σ with unordered Kronecker indices $\kappa_1, \kappa_2, \ldots, \kappa_m$. There is a nice selection ν_κ naturally associated with the unordered Kronecker indices, defined as follows: a (double) index ij belongs to ν_κ iff $i < \kappa_j$. ν_κ is called the Kronecker selection of the pair (F, G). By chasing the definitions, it is easily seen that

(4.4) $\det \rho_{\nu_\kappa}(F, G) \neq 0$.

We will now describe some of the main results of POPOV [1972] on the invariants of the action of $GL(n)$ on Σ. So, let (F, G) be a reachable pair. For the sake of notational simplicity, let rank $G = m$; i. e., (F, G) belongs to $\hat{\Sigma}$. Then Corollary 1 of POPOV [1972] states that there exists a unique set of (indexed) real numbers

$$\{\alpha_{ijk}: i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, i - 1,\ k = 0, 1, \ldots, \min(\kappa_i, \kappa_j - 1);\ \text{or}\ i = 1, 2, \ldots, m,\ j = i, i + 1, \ldots, m,\ k = 0, 1, \ldots, \min(\kappa_i, \kappa_j) - 1\}, \tag{4.5}$$

such that

$$F^{\kappa_i} g_i = \sum_{j=1}^{m} \sum_{k} \alpha_{ijk} F^k g_j, \tag{4.6}$$

where for each i, j the summation with respect to k runs over the index set defined above. $\{\alpha_{ijk}\}$ determines the linear dependence of $F^{\kappa_i} g_i$ on its antecedents. Note that the index set of α_{ijk}'s in the above list depends only upon the unordered Kronecker indices; we denote this index set by A_κ, where $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_m)$. Let N_κ be the cardinality of A_κ. Let $\alpha(F, G)$ denote the (lexicographically) ordered N_κ-tuple (α_{ijk}). We can now state

(4.7) THEOREM (POPOV [1972, THEOREM 1]). *Let* (F_1, G_1) *and* (F_2, G_2) *be in* $\hat{\Sigma}$. *Then* (F_1, G_1) *and* (F_2, G_2) *belong to the same*

orbit under the action χ of $GL(n)$ if and only if (i) the unordered Kronecker indices of (F_1, G_1) and (F_2, G_2) are the same and (ii) $\alpha(F_1, G_1) = \alpha(F_2, G_2)$.

Further, if $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_m)$ is an m-tuple of positive integers satisfying $\kappa_i \geq 1$ and

$$\kappa_1 + \kappa_2 + \ldots + \kappa_m = n,$$

and α is any N_κ-tuple of real numbers, then there exists a pair (F, G) in $\hat{\Sigma}$ such that $\kappa_1, \kappa_2, \ldots, \kappa_m$ are unordered Kronecker indices of (F, G) and $\alpha(F, G) = \alpha$.

Roughly speaking, unordered Kronecker indices and $\alpha(\cdot,\cdot)$ constitute a complete, independent set of invariants for the action of $GL(n)$ on Σ.

We will now describe the algebro-geometric structure of the set of reachable pairs having the same unordered Kronecker indices. Let $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_m)$ be a fixed set of integers satisfying $\kappa_i \geq 1$, and $\kappa_1 + \kappa_2 + \ldots + \kappa_m = n$. Let $\Sigma(\kappa)$ be the subset of Σ defined by $\Sigma(\kappa) := \{(F, G)$ in Σ: the unordered Kronecker indices of (F, G) are $\kappa_1, \kappa_2, \ldots, \kappa_m\}$. We now have

(4.8) THEOREM. Let $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_m)$ be as above. Then $\Sigma(\kappa)$ has the structure of an affine variety. The function $\alpha\colon \Sigma(\kappa) \to \underline{\underline{R}}^{N_\kappa}\colon (F, G) \mapsto \alpha(F, G)$ is a regular surjective map. The pair $(\alpha, \underline{\underline{R}}^{N_\kappa})$ constitutes a geometric quotient of $\Sigma(\kappa)$ under the action of $GL(n)$.

PROOF. We first set up some notation. For each integer i, $1 \leq i \leq m$, and any (F, G) in $\underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m}$, let $\mathrm{Ant}_i(F, G)$ denote the submatrix of $\rho(F, G)$ consisting of the columns indexed by $j\ell$ in the nice selection ν_κ satisfying $\ell + mj < i + m\kappa_i$. In other words, $\mathrm{Ant}_i(F, G)$ consists of antecedents of $F^{\kappa_i} g_i$ which are in $\rho_{\nu_k}(F, G)$. Define the (polynomial) map

$$\mathrm{Ant}_i\colon \underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m} \to \underline{\underline{R}}^{n\times N_i}\colon (F, G) \mapsto \mathrm{Ant}_i(F, G).$$

Here N_i denotes the (fixed) number of columns of $\mathrm{Ant}_i(F, G)$. Let J_κ denote the set of integers j, $1 \leq j \leq m$ such that $N_j < n$.

We will now give equations which give $\Sigma(\kappa)$ as a subset of $\underline{\underline{R}}^{n\times n} \times \underline{\underline{R}}^{n\times m}$. We start by noting that

(4.4) $\quad \det \rho_{\nu_K}(F, G) \neq 0$

is a necessary condition for (F, G) to be in $\Sigma(\kappa)$. The additional condition is that $F^{\kappa_i} g_i$ be linearly dependent upon its antecedents in ν_K; i. e., $F^{\kappa} g_i$ should be linearly dependent upon the columns of $\mathrm{Ant}_i(F, G)$. For any matrix B, let $m_j(B)$ denote the j-th largest minor of B, $j = 1, 2, \ldots, J$. Then we must have

(4.9) $\quad m_j(\mathrm{Ant}_i(F, G): F^{\kappa_i} g_i) = 0, \quad i \text{ in } J_\kappa ,$

for all possible j. It is easily seen that (4.4) and (4.9) are necessary and sufficient in order that $\kappa_1, \kappa_2, \ldots, \kappa_m$ be the unordered Kronecker indices of (F, G).

Following the standard procedure of algebraic geometry, we now introduce a new variable t. Then $\Sigma(\kappa)$ is isomorphic (as a quasi-projective variety) to the (real) solutions of the equations

(4.10) $\quad t \det \rho_{\nu_k}(F, G) = 1$

and the Equations (4.9). This proves our assertion that $\Sigma(\kappa)$ has the structure of an affine variety.

We will now show that the map

$$\alpha: \Sigma(\kappa) \to \underline{\underline{R}}^{N_K}: (F, G) \to \alpha(F, G)$$

is a surjective regular map. Surjectivity follows directly from the results of POPOV [1972] as described in Theorem (4.7). To prove that α is regular, it is enough to show that

(4.11) $\quad \alpha_{ijk}: \Sigma(\kappa) \to \underline{\underline{R}}: (F, G) \mapsto \alpha_{ijk}(F, G)$

is regular for each index ijk in the index set A_K. Now, the map

$$(F, G) \mapsto \det (\rho_{\nu_K}(F, G))^{-1} = t$$

is a regular map on $\Sigma(\kappa)$. It follows that

$$\varphi_{ij}: \Sigma(\kappa) \to \underline{\underline{R}}: (F, G) \mapsto (i, j)\text{-th entry of } (\rho_{\nu_K}(F, G))^{-1},$$

is also a regular map for each i, j. Now for any (F, G) in $\Sigma(\kappa)$, α_{ijk} is (implicitly) given by

$$(4.12) \quad F^{K}i g_i = \sum_{j,k} \alpha_{ijk} F^{k} g_j.$$

Let q denote the position of $F^{j}g_k$ in $\rho_{\nu_K}(F, G)$. Multiplying (4.12) by the q-th row of $(\rho_{\nu_K}(F, G))^{-1}$, we get

$$(\varphi_{q1}, \varphi_{q2} \cdots \varphi_{qn}) F^{K}i g_i = \alpha_{ijk}.$$

As φ_{ij}'s and $F^{K}i g_i$ are regular on $\Sigma(\kappa)$, it follows that α_{ijk} is a regular map. Hence α is a surjective regular map.

Finally, (F_1, G_1) and (F_2, G_2) in $\Sigma(\kappa)$ belong to the same orbit under the action of $GL(n)$ if and only if

$$\alpha(F_1, G_1) = \alpha(F_2, G_2).$$

It follows directly from the definition of geometric quotient, (see MUMFORD [1965],) that $(\alpha, \underline{\underline{R}}^{N_K})$ is a geometric quotient. □

The above result should be viewed as geometric invariant theoretic reformulation of results of POPOV [1972]. We see that $\hat{\Sigma}$ is a disjoint union of affine varieties corresponding to various choices of unordered Kronecker indices; on each of these affine varieties a geometric quotient with respect to the action of $GL(n)$ exists. In our work, we will be working with these affine varieties.

(4.13) REMARK. We note that in the above theorem, <u>existence</u> of the

geometric quotient for the action of $GL(n)$ on $\Sigma(\kappa)$ can be deduced in a much easier way. It is known from TANNENBAUM [1980, Part 3, Section 1, Corollary 1.6] that $GL(n)$ acts on Σ, and hence on $\Sigma(\kappa)$, with closed orbits. Recall that $GL(n)$ is a linearly reductive group. (See FOGARTY [1965].) Having proved that $\Sigma(\kappa)$ is an affine variety, it then follows from MUMFORD [1965, pages 27-30] that a geometric quotient exists. The point here, of course, is that this geometric quotient can be shown to be $(\alpha, \underline{\underline{R}}^{N_\kappa})$.

(4.14) REMARK. Results of the above theorem should be compared with certain results of HAZEWINKEL [1980]. It is shown there that the constructions of the well known "control canonical forms" are, in general, "discontinuous". However, if one restricts attention to these constructions on $\Sigma(\kappa)$, then these constructions are indeed continuous. It is well known that these constructions are closely related to the invariants α_{ijk}. We have shown that α_{ijk}'s are regular maps on $\Sigma(\kappa)$. Thus the continuity of the constructions of canonical forms on $\Sigma(\kappa)$ is natural.

5. Action of the Feedback Group.

We now turn our attention to the action of the feedback group on Σ. This problem has been completely resolved due to the deep and striking results of BRUNOVSKY [1970], ROSENBROCK [1970], KALMAN [1972], POPOV [1972], WONHAM [1974], and others. The main purpose of this section is to describe the stabilizer in the feedback group in a pair (F, G). We start with a theorem that describes a complete set of invariants under the action of the feedback group.

(5.1) THEOREM. The ordered Kronecker indices constitute a complete set of invariants of the action of the feedback group Γ_f on Σ.

For various proofs of this result, the reader is referred to the papers mentioned above.

Let (F, G) be in Σ with ordered Kronecker indices $\kappa_1 \geq \kappa_2 \geq \ldots \geq \kappa_m \geq 1$. Then it is well known that by some element of the feedback group, (F, G) can be transformed to (F_*, G_*);

$$F_* := \text{block diag } [F_1\ F_2\ \dots\ F_m],$$

$$G_* := \text{block diag } [g_1\ g_2\ \dots\ g_m];$$

where F_i, g_i are $\kappa_i \times \kappa_i$, $\kappa_i \times 1$ matrices of the form

$$F_i = \begin{bmatrix} 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & . & . & . & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & 0 & . & . & . & 1 \\ 0 & 0 & 0 & . & . & . & 0 \end{bmatrix}, \quad g_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

The pair (F_*, G_*) is called the Kronecker canonical form of (F, G). In fact, a triple (T, V, L) that transforms (F, G) to (F_*, G_*) can be constructed using the invariants α_{ijk}. In other words, there exists a regular map

$$\delta\colon \Sigma(\kappa) \to \Gamma_f$$

such that $\delta(F, G)$ transforms (F, G) to the Kronecker canonical form.

For our purposes, however, it is more useful to choose a slightly different canonical form for (F, G) under the feedback group. In order to describe this canonical form, we shall first set up some notation. Let $r := \kappa_1 \geq \kappa_2 \geq \dots \geq k_m \geq 1$ be a given set of ordered Kronecker indices. Define the Young's matrix Y_κ associated with this set of indices to be an $m \times n$ matrix given by

$$Y_\kappa(i, j) = \begin{cases} 1, & \text{if } j \leq \kappa_i, \\ 0, & \text{otherwise.} \end{cases}$$

Then the dual Kronecker indices ℓ_j, $1 \leq j \leq r$, are defined by

$$\ell_{r-i+1} = \sum_{j=1}^{m} Y_\kappa(i, j), \quad i = 1, 2, \dots, r.$$

ℓ_j's are the dual indices in the sense of Young's diagrams to the Kronecker indices κ_i's. Let $r(0) = 0 < r(1) < \dots < r(q) := r$

be the unique integers such that

$$\ell_1 = \ell_2 \cdots = \ell_{r(1)} < \ell_{r(1)+1} \cdots = \ell_{r(2)} < \cdots = \ell_{r(q-1)} < \ell_{r(q-1)+1}$$
$$\cdots = \ell_{r(q)} = m.$$

To each integer i, $1 \leq i \leq r$, we can uniquely associate another integer $\beta(i)$ defined by $\ell_i = \ell_{r(\beta(i))}$. Finally, define

$$\sigma_1 = \ell_{r(1)}, \quad \sigma_2 = \ell_{r(2)} - \ell_{r(1)}, \quad \cdots, \quad \sigma_q = \ell_{r(q)} - \ell_{r(q-1)}.$$

We note that in case the Kronecker indices are all equal; i. e., $\kappa_1 = \kappa_2 = \cdots = \kappa_m = n/m$, then

$$\ell_1 = \ell_2 = \cdots = \ell_{\kappa_1} = m, \quad q = 1, \quad \sigma_1 = m, \quad r(1) = \kappa_1.$$

As we will see in the following, the case of equal Kronecker indices is relatively uncomplicated and leads to very elegant results.

Now let (F, G) be in Σ with ordered Kronecker indices $\kappa_1 \geq \kappa_2 \geq \cdots \geq \kappa_m \geq 1$. Then it is easy to see that (F, G) is feedback equivalent to (F^*, G^*);

$$(5.2) \qquad F^* = \begin{bmatrix} 0 & F_1 & 0 & \cdots & 0 \\ 0 & 0 & F_2 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & F_{r-1} \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad G^* = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_m \end{bmatrix}.$$

Here "0" in ij-th entry of F^* is the $\ell_i \times \ell_j$ matrix consisting of zeroes, F_i is the $\ell_i \times \ell_{i+1}$ matrix of the form

$$F_i = [I_{\ell_i} \; 0],$$

and "0" in the i-th row of G^* denotes the $\ell_i \times \ell_{i+1}$ matrix

consisting of zeroes. In case the Kronecker indices are equal, we have

$$F_1 = F_2 = \cdots = F_{r-1} = I_m.$$

We call (F^*, G^*) the dual Kronecker canonical form. It is easy to see that (F^*, G^*) can be obtained from (F_*, G_*) by a permutation of state variables. Here again, there exists a regular map, ("built" out of α_{ijk}'s),

$$d: \Sigma(\kappa) \to \Gamma_f: (F, G) \mapsto d(F, G),$$

such that $d(F, G)$ transforms (F, G) to (F^*, G^*). Throughout this dissertation, (F^*, G^*) will denote the dual Kronecker canonical form for some choice of Kronecker indices. We now turn our attention to the stabilizer in the feedback group of a pair (F^*, G^*). Stabilizer in the feedback group of the pair (F_*, G_*) (in the Kronecker canonical form) has been obtained by WANG and DAVISON [1976], BROCKETT [1977], FUHRMANN and WILLEMS [1979]. For the Riccati group problem, however, it turns out that the stabilizer of (F^*, G^*) (in the dual Kronecker canonical form) is more convenient. Let (T^*, V^*, L^*) be in the feedback group. We partition T^*, V^*, L^* as

$$T^* = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1r} \\ T_{21} & T_{22} & \cdots & T_{2r} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ T_{r1} & T_{r2} & \cdots & T_{rr} \end{bmatrix},$$

$$V^* = \begin{bmatrix} V_{11} & V_{12} & \cdot & \cdot & \cdot & V_{1q} \\ V_{21} & V_{22} & \cdot & \cdot & \cdot & V_{2q} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ V_{q1} & V_{q2} & \cdot & \cdot & \cdot & V_{qq} \end{bmatrix},$$

$$L^* = \begin{bmatrix} L_1 & L_2 & \cdot & \cdot & \cdot & L_r \end{bmatrix},$$

where T_{ij} is an $\ell_i \times \ell_j$, V_{ij} is an $\sigma_i \times \sigma_j$, and L_i is an $m \times \ell_i$ matrix. The stabilizer subgroup of (F^*, G^*) is characterized in the following

(5.3) THEOREM. Let (F^*, G^*) be in the dual Kronecker canonical form as described above. Then a triple (T^*, V^*, L^*) belongs to the stabilizer of (F^*, G^*) if and only if (T^*, V^*, L^*) satisfies:

$$V_{ij} = 0, \quad j > i,$$

$$\det V_{ii} \neq 0, \quad i = 1, 2, \ldots, q,$$

$$T_{ij} = 0, \quad j > i,$$

$$T_{ii} = \begin{bmatrix} V_{11} & V_{12} & \cdot & \cdot & \cdot & V_{1\beta} \\ V_{21} & V_{22} & \cdot & \cdot & \cdot & V_{2\beta} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ V_{\beta 1} & V_{\beta 2} & \cdot & \cdot & \cdot & V_{\beta\beta} \end{bmatrix}, \quad \text{where} \quad \beta = \beta(i);$$

$$T_{ri} = \begin{bmatrix} t^i_{11} & t^i_{12} & \cdots & t^i_{1\beta} \\ t^i_{21} & t^i_{22} & \cdots & t^i_{2\beta} \\ \vdots & & & \\ t^i_{q1} & t^i_{q2} & \cdots & t^i_{q\beta} \end{bmatrix}, \quad \text{where} \quad i < r, \quad \beta = \beta(i),$$

where t^i_{jk} is a $\sigma_j \times \sigma_k$ matrix satisfying

$$t^i_{jk} = 0 \quad \text{for} \quad j \leq \beta(r - 1 + r(k - 1)),$$

$$T_{ij} = \begin{bmatrix} t^\ell_{11} & t^\ell_{12} & \cdots & t^\ell_{1\beta} \\ t^\ell_{21} & t^\ell_{22} & \cdots & t^\ell_{2\beta} \\ \vdots & & & \vdots \\ t^\ell_{\alpha 1} & t^\ell_{\alpha 2} & \cdots & t^\ell_{\alpha\beta} \end{bmatrix}, \quad r > i > j,$$

where

$$\ell = r - i + j < r, \quad \beta = \beta(j), \quad \alpha = \beta(i);$$

finally,

$$L_1 = 0, \quad L_i = T_{r(i-1)}(I_{\ell_{i-1}} \; 0), \quad i = 2, \ldots, r.$$

The proof of this result is omitted since it is a very tedious and cumbersome but straightforward direct verification. We will now state a very simple corollary of the above theorem which describes the stabilizer in case all Kronecker indices are equal. The description in this case becomes simple and manageable.

(5.4) COROLLARY. Let (F^*, G^*) be in the dual Kronecker canonical form with Kronecker indices $\kappa_1 = \kappa_2 = \ldots = \kappa_m = n/m$. Then a triple (T^*, V^*, L^*) belongs to the stabilizer of (F^*, G^*) if and only if

$$T^* = \begin{bmatrix} V & 0 & . & . & . & 0 \\ 0 & V & . & . & . & 0 \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 0 & 0 & . & . & . & V \end{bmatrix}, \qquad V^* = V,$$

$$L^* = 0, \quad \det V \neq 0.$$

The above corollary is much easier to verify directly. In the following chapters, we shall make extensive use of these results. Corollary (5.4) will be especially useful since in this case the results are easy to formulate in an explicit form.

CHAPTER IV. ACTION OF THE RICCATI GROUP--I

In this chapter we shall develop an approach which leads to a complete set of invariants and canonical forms for the action of the Riccati group on certain subsets of the "optimization data space" D. The results of this chapter are most satisfactory for the case of equal Kronecker indices. In other cases the results are similar; however it seems to be difficult to formulate these results in closed form due to notational and computational complexity. An alternative approach which also runs into somewhat analogous difficulties is developed in the next chapter. Some of the results of this chapter will be used in Chapter VI where we discuss system theoretic applications and interpretations of the action of the Riccati transformation group.

This chapter is divided into four parts. First we discuss the finite-time problems, next we give an expository account of the invariant theory of q-tuples of square matrices under the orthogonal group, we then reconsider finite time problems, and finally we consider the infinite horizon-zero final state problems.

6. Finite Time Problems--I.

We recall that the Riccati group Γ acts on the data space D by

$$\varphi\colon D \times \Gamma \to D\colon ((F, G, Q, R, S, P_o), (T, V, L, M)) \mapsto (\hat{F}, \hat{G}, \hat{R}, \hat{S}, \hat{P}_o),$$

where

$$\hat{F} = T^{-1}FT + T^{-1}GL,$$

$$\hat{G} = T^{-1}GV,$$

$$\hat{Q} = T'(Q + F'M + MF)T + T'(S + MG)L + L'(S + MG)'T + L'RL,$$

$$\hat{R} = V'RV,$$

$$\hat{S} = (T'(S + MG) + L'R)V,$$

$$\hat{P}_o = T'(P_o - M)T.$$

We start with the following

(6.1) PROPOSITION. Let (F, G, Q, R, S, P_o) be in D and let the ordered Kronecker indices of (F, G) be $\kappa_1 \geq \kappa_2 \geq \ldots \geq \kappa_m \geq 1$. Then there exists a quadruple (T, V, L, M) in the Riccati group Γ such that $\varphi((F, G, Q, R, S, P_o), (T, V, L, M)) = (F^*, G^*, \hat{Q}, \hat{R}, \hat{S}, 0)$ for some $\hat{Q}, \hat{R}, \hat{S}$. Further, any two six-tuples $(F_1^*, G_1^*, Q_1, R_1, S_1, 0)$ and $(F_2^*, G_2^*, Q_2, R_2, S_2, 0)$ in D belong to the same orbit if and only if the ordered Kronecker indices of (F_1^*, G_1^*) and (F_2^*, G_2^*) are equal, (and hence $F_1^* = F_2^* =: F^*$, $G_1^* = G_2^* =: G^*$,) and there exists a triple (T^*, V^*, L^*) in the stabilizer in the feedback group of (F^*, G^*) such that

$$Q_2 = (T^*)'Q_1T^* + (T^*)'S_1L^* + (L^*)'S_1'T^* + (L^*)'R_1L^*, \tag{6.2}$$

$$S_2 = (T^*)'S_1V^* + (L^*)'R_1V^*, \tag{6.3}$$

$$R_2 = (V^*)'R_1V^*. \tag{6.4}$$

[Here (F^*, G^*) is the dual Kronecker canonical form as explained in Chapter 3.]

PROOF. By the techniques of Chapter III, there exists a triple (T, V, L) such that

$$F^* = T^{-1}FT + T^{-1}GL,$$

$$G^* = T^{-1}GV.$$

Choose $M = P_o$. Then

$$\varphi((F, G, Q, R, S, P_o), (T, V, L, M)) = (F^*, G^*, \hat{Q}, \hat{R}, \hat{S}, 0),$$

where $\hat{Q}, \hat{R}, \hat{S}$ are given by the formulas described at the beginning of this chapter.

Suppose there exists a quadruple (T, V, L, M) such that

$$\varphi((F_1^*, G_1^*, Q_1, R_1, S_1, 0), (T, V, L, M)) = (F_2^*, G_2^*, Q_2, R_2, S_2, 0).$$

Then

$$F_2^* = T^{-1}F_1^*T + T^{-1}G_1^*L, \quad G_2^* = T^{-1}G_1^*V.$$

Hence, by feedback invariance, the Kronecker indices of (F_2^*, G_2^*) must be the same as those of (F_1^*, G_1^*). As (F_1^*, G_1^*) and (F_2^*, G_2^*) are in dual Kronecker canonical form, it follows that $F_1^* = F_2^* =: F^*$ and $G_1^* = G_2^* =: G^*$. Hence, (T, V, L) belongs to the stabilizer in the feedback group of (F^*, G^*). Further,

$$0 = T'(0 + M)T = T'MT.$$

Since T is nonsingular, $M = 0$. Hence, $(T, V, L, M) = (T^*, V^*, L^*, 0)$, where (T^*, V^*, L^*) belongs to the stabilizer of (F^*, G^*). Now the Equations (6.2), (6.3), and (6.4) follow immediately from the definition of φ.

Conversely, if (F_1^*, G_1^*) and (F_2^*, G_2^*) have the same ordered Kronecker indices, and since (F_1^*, G_1^*) and (F_2^*, G_2^*) are in the dual Kronecker canonical form, it follows that $F_1^* = F_2^* =: F^*$, $G_1^* = G_2^* = G^*$. Further, if there exists (T^*, V^*, L^*) in the stabilizer in the feedback group of (F^*, G^*) such that Equations (6.2), (6.3), and (6.4) hold, then it follows directly from the definition of φ that

$$\varphi((F_1^*, G_1^*, Q_1, R_1, S_1, 0), (T^*, V^*, L^*, 0)) = (F_2^*, G_2^*, Q_2, R_2, S_2, 0).$$

Hence, $(F_1^*, G_1^*, Q_1, R_1, S_1, 0)$ and $(F_2^*, G_2^*, Q_2, R_2, S_2, 0)$ belong to the same orbit. □

This proposition shows that the problem of the action of the Riccati group is equivalent to the problem of the action of stabilizer in the feedback group on a certain space of matrices. As we saw in Chapter III, the stabilizer in the feedback group of (F^*, G^*) in the dual Kronecker canonical form is completely determined by the ordered Kronecker indices.

Thus, it is natural to study this problem for various choices of Kronecker indices. The simplest among these cases is the case of equal Kronecker indices.

Let us define a subset D_1 of the optimization problem data space D as

$$D_1 := \{(F, G, Q, R, S, P_o) \text{ in } D : \text{the ordered Kronecker indices of } (F, G) \text{ are equal}\}.$$

It is obvious that D_1 is stable under the action of the Riccati group.

Let (F, G, Q, R, S, P_o) be in D_1 such that the Kronecker indices of (F, G) are $r = \kappa_1 = \kappa_2 = \ldots = \kappa_m = n/m$. By Proposition (6.1) there exists (T, V, L, M) in Γ such that

$$\varphi((F, G, Q, R, S, P_o), (T, V, L, M)) = (F^*, G^*, \hat{Q}, \hat{R}, \hat{S}, 0).$$

Further, we need to consider the action of the stabilizer of the feedback group of (F^*, G^*) on matrices $\hat{Q}$, $\hat{R}$, $\hat{S}$. For any $n \times n$ matrix A and any $n \times m$ matrix B, let A_{ij}, B_k be the $m \times m$ matrices such that

$$A = \begin{bmatrix} A_{11} & A_{22} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & & & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rr} \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_r \end{bmatrix}.$$

With this notation, we now have

(6.5) PROPOSITION. *Let (F, G, Q, R, S, P_o) in D be such that the Kronecker indices of (F, G) are $\kappa_1 = \kappa_2 = \ldots = \kappa_m = n/m$. Then there exists a quadruple (T, V, L, M) in Γ such that*

$$\varphi((F, G, Q, R, S, P_o), (T, V, L, M)) = (F^*, G^*, \hat{Q}, I, \hat{S}, 0).$$

Further, any two such six tuples $(F^*, G^*, Q^1, I, S^1, 0)$ and $(F^*, G^*, Q^2, I, S^2, 0)$ with equal Kronecker indices belong to the same orbit if and only if there exists an $m \times m$ orthogonal matrix V such that

$$(6.6) \qquad V'Q^1_{ij}V = Q^2_{ij}, \quad V'S^1_kV = S^2_k, \quad i, j, k = 1, 2, \ldots, r.$$

PROOF. By Proposition (6.1), there exists a quadruple (T, V, L, M) in Γ such that

$$\varphi((F, G, Q, R, S, P_o), (T, V, L, M)) = (F^*, G^*, \tilde{Q}, \tilde{R}, \tilde{S}, 0).$$

Let $\tilde{V}$ be an $m \times m$ nonsingular matrix such that $\tilde{V}'\tilde{R}\tilde{V} = I$. Such a $\tilde{V}$ exists since $\tilde{R}$ is a positive definite symmetric matrix. Define $\tilde{T} := (\tilde{V}, \tilde{V}, \ldots, \tilde{V})$. Then it is easy to see that

$$\varphi((F^*, G^*, \tilde{Q}, \tilde{R}, \tilde{S}, 0), (\tilde{T}, \tilde{V}, 0, 0)) = (F^*, G^*, \hat{Q}, I, \hat{S}, 0).$$

Suppose there exists a (T, V, L, M) in Γ such that

$$\varphi((F^*, G^*, Q^1, I, S^1, 0), (T, V, L, M)) = (F^*, G^*, Q^2, I, S^2, 0).$$

Then by Proposition (6.1), $M = 0$ and (T, V, L) belongs to the stabilizer of the feedback group of (F^*, G^*). Hence $L = 0$. Further, $V'V = I$ implies that V is orthogonal. By Corollary (5.4), $T = \text{diag}\,(V, V, \ldots, V)$. Equations (6.6) follow directly from Equations (6.2) and (6.3).

Conversely, if V is any $m \times m$ orthogonal matrix such that (6.6) holds, then it follows directly from the definition of φ that

$$\varphi((F^*, G^*, Q^1, I, S^1, 0), (T, V, 0, 0)) = (F^*, G^*, Q^2, I, S^2, 0).$$

where $T := \text{diag}\,(V, V, \ldots, V)$. This completes the proof. □

This proposition shows that the problem of the action of the Riccati group on "data space" corresponding to equal Kronecker indices can be

reduced to the action of the orthogonal group on matrices Q_{ij}, S_k as given by (6.6). It will be seen later that in some other cases as well, the problem is essentially that of the action of the orthogonal group on a bunch of matrices. In view of this, we make a brief digression and describe the invariant theory of q-tuples of matrices under the action of the orthogonal group.

7. The Orthogonal Group

The purpose of this section is to describe certain somewhat recent results on the invariants of the action of the orthogonal group on a space of matrices. Most of these results are originally due to SIBIRSKII [1967], which were later rediscovered in an extended form by PROCESI [1976]. We will state only those results which are relevant to our problems. For detailed proofs and related results involving other classical groups, see SIBIRISKII [1967] and PROCESI [1976].

Let W be the space of q-tuples of $m \times m$ matrices. The group $O(m)$ of $m \times m$ (real) orthogonal matrices acts on W as given by

$$\chi: W \times O(m) \to W:((A_1, A_2, \ldots, A_q), V) \mapsto (V'A_1V, V'A_2V, \ldots, V'A_qV).$$

The problem is to find invariants of the action χ. Let $\underline{\underline{R}}[W]$ denote the ring of real valued polynomial functions on W. It is easy to see that $\underline{\underline{R}}[W]$ is isomorphic to the ring of polynomials in qm^2 indeterminates with real coefficients. Let RO be the subring of $\underline{\underline{R}}[W]$ consisting of those functions which remain invariant under the action χ of $O(m)$ on W, i. e., f in $\underline{\underline{R}}[W]$ belongs to RO if and only if

$$f((A_1, A_2, \ldots, A_q)) = f((V'A_1V, V'A_2V, \ldots, V'A_qV))$$

for all $(A_1, A_2, \ldots, A_q)$ in W and V in $O(m)$. Let $\alpha_1, \alpha_2, \ldots, \alpha_q, \beta_1, \beta_2, \ldots, \beta_q$ be noncommuting indeterminates. For any monomial $f(\alpha_1, \alpha_2, \ldots, \alpha_q, \beta_1, \beta_2, \ldots, \beta_q)$ in these indeterminates with real coefficients, we associate a polynomial

function

$$\text{tr } f\colon W \to \underline{\underline{R}}\colon (A_1, A_2, \ldots, A_q) \mapsto \text{tr}(f(A_1, A_2, \ldots, A_q, A_1', A_2', \ldots, A_q')).$$

We now have the main

(7.1) THEOREM. (SIBIRSKII [1967, Lemma 2], PROCESI [1976, Theorem 7.3]). The ring RO is generated as an $\underline{\underline{R}}$-algebra by the elements tr f, where f is a monomial in noncommutative indeterminates $\alpha_1, \alpha_2, \ldots, \alpha_q, \beta_1, \beta_2, \ldots, \beta_q$ with $\deg f \leq 2^m - 1$. Further, any two q-tuples $(A_1, A_2, \ldots, A_q)$ and $(B_1, B_2, \ldots, B_q)$ belong to the same orbit if and only if

$$\text{tr } f\,((A_1, A_2, \ldots, A_q)) = \text{tr } f\,((B_1, B_2, \ldots, B_q))$$

for all monomials f in the above mentioned generating set.

Let $f_1, f_2, \ldots, f_N$ be all the distinct monomials in the noncommutative variables $\alpha_1, \alpha_2, \ldots, \alpha_q, \beta_1, \beta_2, \ldots, \beta_q$ of degree $\leq 2^m - 1$. For a given q-tuple $A = (A_1, A_2, \ldots, A_q)$ in W, let

$$\text{Tr } A := (\text{tr } f_1(A), \text{ tr } f_2(A), \ldots, \text{tr } f_N(A)).$$

Then, in light of the above theorem, Tr A is a complete invariant for the action of O(m) on W.

It is known that the orthogonal group is a linearly reductive group. Thus, χ is the action of a linearly reductive group on an affine variety. By Theorem (5.9) of FOGARTY [1969], it follows immediately that the ring of invariants RO is finitely generated. Theorem (7.1) gives explicit formulae for a set of generators.

We now consider the problem of canonical forms for the action χ of O(m) on W. In a private communication, C. Procesi has indicated that it is likely that global canonical forms do not exist. However, we have been able to get canonical forms on certain dense open

subsets (in the Euclidean topology) of W. For any integer i, $1 \leq i \leq q$, define

$$W_i = \{A = (A_1, A_2, \ldots, A_q) \text{ in } W: \text{ the (real) eigenvalues of } (A_i + A_i')\ \text{are distinct}\}.$$

Now, let A be in W_i. Let V be an orthogonal matrix such that

$$V'(A_i + A_i')V/2 = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n), \quad \lambda_1 > \lambda_2 > \ldots > \lambda_n.$$

Further

$$V'A_iV = V'(A_i + A_i')V/2 + V'(A_i - A_i')V/2$$

$$= \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n) + \text{a skew-symmetric matrix.}$$

We call a matrix M quasi-skew-symmetric iff M is the sum of a diagonal matrix (with decreasing entries) and a skew-symmetric matrix. Thus, any A in W_i can be transformed by an orthogonal matrix to $\hat{A}$ such that $\hat{A}_i$ is quasi-skew-symmetric.

Now let A, B in W_i be such that A_i, B_i are quasi-skew-symmetric. If A, B belong to the same orbit under $O(m)$, then there exists an orthogonal matrix V such that $V'A_jV = B_j$, $j = 1, 2, \ldots, q$. In particular,

$$V'A_iV = B_i.$$

Since A_i, B_i are quasi-skew-symmetric, it follows that the diagonal parts of A_i and B_i are equal. Since the diagonal entries of A_i and B_i are distinct, it then follows from elementary linear algebra that

$$V = \text{diag}\,(\epsilon_1, \epsilon_2, \ldots, \epsilon_m), \quad \epsilon_i = +1 \text{ or } -1.$$

This ambiguity can be resolved by fixing signs of certain (non zero) off diagonal entries of A_j's. Thus, on each W_i we can obtain a canonical form under the action of χ.

Finally, we note that each W_i is a dense open subset of W. For, A belongs to W_i if and only if the characteristic polynomial of $(A_i + A_i')$ has no repeated roots. Using resultants, it is easy to show that W_i is the complement of an algebraic set of strictly lower dimension.

8. Finite Time Problems--II

We now resume our discussion of Section 6. It was shown in Proposition (6.5) that for the case of equal Kronecker indices, every orbit contains a six-tuple of the form $(F^*, G^*, Q, I, S, 0)$. Let (Q, S) denote the $((n/m)^2 + (n/m)) =: q$-tuple consisting of the $m \times m$ matrices Q_{ij}, S_k ordered in some fixed way. It now follows immediately from Proposition (6.5) and Theorem (7.1) that $Tr(Q, S)$ forms a complete invariant. More precisely, we have

(8.1) THEOREM. *Let* (F, G, Q, R, S, P_o) *be in* D_1. *Then there exist* $\hat{Q}, \hat{S}$ *such that* $(F^*, G^*, \hat{Q}, I, \hat{S}, 0)$ *is in the same orbit as* (F, G, Q, R, S, P_o) *under the action of the Riccati group. Let* $TR[(F, G, Q, R, S, P_o)]$ *denote* $Tr(\hat{Q}, \hat{S})$. *Then*

$$TR: D_1 \to \underline{\underline{R}}^N: (F, G, Q, R, S, P_o) \mapsto TR[(F, G, Q, R, S, P_o)]$$

is a well defined function. Further, any two six-tuples $d_1 = (F_1, G_1, Q_1, R_1, S_1, P_{o_1})$ *and* $d_2 = (F_2, G_2, Q_2, R_2, S_2, P_{o_2})$ *in* D_1 *belong to the same orbit if and only if* $TR[d_1] = TR[d_2]$.

PROOF. The theorem follows immediately from Propositions (6.5) and (7.1). □

We note that using the techniques outlined at the end of the previous section, we can obtain canonical forms for the action of the Riccati group.

We will now treat another special case which corresponds to strictly unequal Kronecker indices. Let us define

$$D_2 := \{(F, G, Q, R, S, P_o): \text{ the ordered Kronecker indices of } (F, G) \text{ are } \kappa_1 > \kappa_2 > \ldots > \kappa_m \geq 1\}.$$

In this case, we shall outline a procedure that leads to a canonical form (and hence a complete set of invariants) for the action of the Riccati group on D_2. In order to motivate this procedure, we start with an example.

Let $\kappa_1 = 2 > \kappa_2 = 1$. The dual Kronecker canonical form is

$$(8.2) \qquad F^* = \left[\begin{array}{c|cc} 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \qquad G^* = \left[\begin{array}{cc} 0 & 0 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right]$$

Now (T^*, V^*, L^*) belongs to the stabilizer of the feedback group of (F^*, G^*) if and only if

$$T^* = \left[\begin{array}{c|cc} v_{11} & 0 & 0 \\ \hline 0 & v_{11} & 0 \\ t & v_{21} & v_{22} \end{array}\right], \qquad V^* = \left[\begin{array}{cc} v_{11} & 0 \\ v_{21} & v_{22} \end{array}\right],$$

$$(8.3)$$

$$L^* = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & t & 0 \end{array}\right], \qquad v_{11} \neq 0, \quad v_{22} \neq 0, \quad v_{21}, \quad t \text{ arbitrary.}$$

Now by Proposition (6.1), we can restrict our attention to six-tuples of the form $(F^*, G^*, \hat{Q}, \hat{R}, \hat{S}, 0)$. For any given R, there exists a lower triangular matrix V^* such that

$$(V^*)'RV^* = I.$$

Then, $(T^*, V^*, 0, 0)$ transforms $(F^*, G^*, \hat{Q}, \hat{R}, \hat{S}, 0)$ to $(F^*, G^*, Q, I, S, 0)$. Now let

$$V^* = I, \quad L^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & t & 0 \end{bmatrix}, \quad T^* = \begin{bmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then under the action of $(T^*, V^*, L^*, 0)$, $(F^*, G^*, Q, I, S, 0)$ transforms to $(F^*, G^*, Q^1, I, S^1, 0)$ where

$$Q^1 = (T^*)'QT^* + (T^*)'SL^* + (L^*)'ST + (L^*)'L, \quad S^1 = (T^*)'S + (L^*)'.$$

Let us consider the transformation that S undergoes: (s_{ij} denotes ij-th entry of S),

$$s_{11} \to s_{11} + s_{31}, \quad s_{12} \to s_{12} + ts_{32},$$

$$s_{21} \to s_{21}, \quad s_{22} \to s_{22} + t,$$

$$s_{31} \to s_{31}, \quad s_{32} \to s_{32}.$$

Choose $t = -s_{22}$. Then $s_{22} \to 0$. This shows that any element of D_2 can be transformed to $(F^*, G^*, Q, I, S, 0)$ such that $s_{22} = 0$.

Further, $(F^*, G^*, Q^1, I, S^1, 0)$ and $(F^*, G^*, Q^2, I, S^2, 0)$ (with $s^1_{22} = 0$, $s^2_{22} = 0$) belong to the same orbit if and only if there exist $\epsilon_i = +1$ or -1, and

$$\begin{bmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_2 \end{bmatrix} Q^1 \begin{bmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_2 \end{bmatrix} = Q^2,$$

$$\begin{bmatrix} \epsilon_2 & 0 & 0 \\ 0 & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_2 \end{bmatrix} S^1 \begin{bmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{bmatrix} = S^2.$$

In other words, $(F^*, G^*, Q, I, S, 0)$ with $s_{22} = 0$ is a <u>canonical form</u> up to the action of "signature matrices" as described above.

Consequently, modulo the action of such signature matrices, the entries of Q, S in the above form constitute a complete set of invariants.

In general, we start by noting that any element of D_2 can be brought into the form $(F^*, G^*, Q, R, S, 0)$ by a suitable element of the Riccati group. Here (F^*, G^*) denotes the dual Kronecker canonical form associated with Kronecker indices $\kappa_1 > \kappa_2 > \ldots > \kappa_m \geq 1$. Specializing Theorem (5.3) to this case, it is easy to see that (T^*, V^*, L^*) belongs to the stabilizer in the feedback group of (F^*, G^*) if and only if

$$T^* = \begin{bmatrix} T_{11} & 0 & \cdot & \cdot & \cdot & 0 \\ T_{21} & T_{22} & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ T_{r1} & T_{r2} & \cdot & \cdot & \cdot & T_{rr} \end{bmatrix}, \quad V^* = \begin{bmatrix} v_{11} & 0 & \cdot & \cdot & \cdot & 0 \\ v_{21} & v_{22} & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ v_{m1} & \cdot & \cdot & \cdot & \cdot & v_{mm} \end{bmatrix},$$

$$L^* = [0 \quad L_2 \quad L_3 \quad \cdot \quad \cdot \quad \cdot \quad L_r], \quad v_{ii} \neq 0, \quad i = 1, 2, \ldots, m,$$

and T_{ij}, L_i, v_{ij}'s satisfy the formulae in Theorem (5.3). Since R is positive definite, there exists an upper triangular matrix V^* such that

$$(V^*)'RV^* = I.$$

Then, with $L^* = 0$ and the appropriate T^* such that $(T^*, V^*, 0)$ is in the stabilizer of (F^*, G^*), $(T^*, V^*, 0, 0)$ transforms $(F^*, G^*, Q, R, S, 0)$ to $(F^*, G^*, \hat{Q}, I, \hat{S}, 0)$ for some $\hat{Q}$, $\hat{S}$. Furthermore, two elements $(F^*, G^*, Q^1, I, S^1, 0)$ and $(F^*, G^*, Q^2, I, S^2, 0)$ belong to the same orbit under Γ if and only if there exists (T^*, V^*, L^*) in the stabilizer in the feedback group of (F^*, G^*) such that

$$(V^*)'V^* = I$$

and (6.2) and (6.3) hold. But as V^* is lower triangular, this implies that $V^* = \text{diag}(\epsilon_1, \epsilon_2, \ldots, \epsilon_m)$, $\epsilon_i = +1$ or -1. Let us consider the action of such elements of the feedback group on S as given by (6.3). Partition S as

$$S' = [S_1', S_2', \ldots, S_r'].$$

Choose $V^* = I_m$. It is then easy to verify that $T_{ii} = I_{\ell_i}$. Under the action of (T^*, I, L^*),

$$S_i \to S_i + T'_{(i+1)i} S_{i+1} + \cdots + T'_{ri} S_r + L_i',$$

$$i = 1, 2, \ldots, r.$$

Let A_L be the set of indices ij such that the ij-th entry of $(L^*)'$ is free. Then we can recursively solve for L^* such that the (ij)-th entry of S is zero for ij in A_L. We start with

$$S_r \to S_r + L_r.$$

It is clear that L_r can be chosen so that the entries of S_r corresponding to indices in A_L are zero. It can be easily verified that this choice of L_r fixes $T_{r(r-1)}, T_{(r-1)(r-2)}, \cdots, T_{21}$. Next we consider

$$S_{r-1} \mapsto S_{r-1} + T'_{r(r-1)} S_r + L_{r-1}.$$

Since $T_{r(r-1)}$ is fixed, we can find L_{r-1} to trivialize certain entries of S_{r-1} which correspond to indices A_L. Clearly, this procedure can be recursively completed to obtain the desired L^* and T^*.

Hence, we may assume that $(F^*, G^*, Q, I, S, 0)$ is in such a form that the entries of S corresponding to indices in A_L are zero. Let us denote the set of such matrices by SC. It is straightforward, though tedious, to verify that $(F^*, G^*, Q^1, I, S^1, 0)$ and $(F^*, G^*, Q^2, I, S^2, 0)$, with S^1, S^2 in SC, are in the same orbit if and only if there exist $\epsilon_1, \epsilon_2, \ldots, \epsilon_m$, $\epsilon_i = +1$ or -1, such that with

$$V^* = \text{diag}\,(\epsilon_1, \epsilon_2, \ldots, \epsilon_m),$$

$$T^*_{ii} = \text{diag}\,(\epsilon_1, \epsilon_2, \ldots, \epsilon_{\ell_i}),$$

$$T^*_{ij} = 0, \quad \text{for} \quad i \neq j,$$

we have

(8.4) $\quad T^* Q^1 V^* = Q^2, \quad T^* S^1 V^* = S^2.$

In other words, $(F^*, G^*, Q, I, S, 0)$ with S in SC is a canonical form modulo the action of signature matrices as described in (8.4).

We now define an equivalence relation of pairs of matrices (Q, S). Any two pairs of matrices (Q^1, S^1) and (Q^2, S^2) with Q^1, Q^2 $n \times n$ symmetric matrices and S^1, S^2 in SC are said to be <u>sign equivalent</u> if there exist $\epsilon_1, \epsilon_2, \ldots, \epsilon_m$, $\epsilon_i = +1$ or -1, such that Equations (8.4) hold. Clearly, sign equivalence is an equivalence relation. Let $[Q, S]$ denote the equivalence class of the pair (Q, S). Let QS denote the collection of these equivalence classes. It is obvious that each $[Q, S]$ consists of a <u>finite collection of pairs of matrices</u>. We can now state the main

(8.5) THEOREM. <u>Let</u> (F, G, Q, R, S, P_o) <u>be any element of</u> D_2. <u>Then there exist matrices</u> $\hat{Q}, \hat{S}$ <u>with</u> $\hat{S}$ <u>in</u> SC <u>such that</u> $(F^*, G^*, \hat{Q}, I, \hat{S}, 0)$ <u>is in the same orbit as</u> (F, G, Q, R, S, P_o) <u>under the action of the Riccati group. Further, the assignment</u>

$$qs: D_2 \to QS: (F, G, Q, R, S, P_o) \mapsto [\hat{Q}, \hat{S}]$$

<u>is a well defined function from</u> D_2 <u>onto</u> QS. <u>Finally, for any</u> d_1, d_2 <u>in</u> D_2, $qs(d_1) = qs(d_2)$ <u>if and only if</u> d_1 <u>and</u> d_2 <u>belong to the same orbit under the Riccati Group</u>.

The above theorem gives a complete set of invariants in terms of finite collections of pairs of matrices which can also be used to construct canonical forms. We note the function qs which assigns invariants is <u>not</u> a polynomial function of the data. This is mainly due to the fact that in order to arrive at the canonical form, we had to find a square root of R.

Thus, we see that Theorems (8.1) and (8.5) solve the problem of invariants and canonical forms for the cases of equal Kronecker indices and distinct Kronecker indices. The general case is a "combination" of these two cases. Analysis of the general case requires combining the action of the orthogonal group and the "affine action" of L. However, it seems very difficult to formulate any general results along this direction. In the next chapter, we shall consider an entirely different approach which also leads to invariants and canonical forms, but runs into analogous difficulties.

9. Infinite Time-Zero Terminal State Case--I

As discussed in Chapter II, this problem arises when we consider the action of the Riccati group on quintuples (F, G, Q, R, S). In this section, we restrict our attention to the subset of data space which corresponds to optimization problems with equal Kronecker indices. We have not been able to extend the techniques of this section to the general case to obtain invariants/canonical forms. In the next chapter, we shall reexamine this problem from an entirely different point of view which leads to invariants and canonical forms.

Let us define DS_1 to be the subset of DS as $DS_1 := \{(F, G, Q, R, S)$ in DS: the ordered Kronecker indices of (F, G) are equal$\}$. We start by choosing a Hurwitz polynomial

$$\chi(s) = s^r - \alpha_1 s^{r-1} - \alpha_2 s^{r-2} - \ldots - \alpha_r.$$

(For example, choose $\chi(s) = (s + 1)^r$.) Given a reachable pair (F, G) with Kronecker indices $\kappa_1 = \kappa_2 = \ldots = \kappa_m = n/m = r$, there exist a triple (T, V, L) in the feedback group such that

$$T^{-1}FT + T^{-1}GL = F_s^*, \quad T^{-1}GV = G^*,$$

where F_s^* is given by

$$F_s^* = \begin{bmatrix} 0 & I_m & 0 & \cdots & 0 \\ 0 & 0 & I_m & & 0 \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & I_m \\ \alpha_r I_m & \alpha_{r-1} I_m & \cdots & & \alpha_1 I_m \end{bmatrix} =: F^* + G^* L_\alpha .$$

It is easy to verify that the characteristic polynomial of F_s^* is χ^m and hence the eigenvalues of F_s^* are in the left half plane. We now have a preliminary

(9.1) LEMMA. A triple (T, V, L) belongs to the stabilizer of the feedback group of (F_s^*, G^*) if and only if

V = a nonsingular $m \times m$ matrix,

$T = \operatorname{diag}(V, V, \ldots, V)$, $L = 0$.

PROOF. If $T^{-1}F_s^*T + T^{-1}GL = F_s^*$ and $T^{-1}GV = G^*$, then

$$T^{-1}F^*T + T^{-1}G^*L_\alpha T + T^{-1}G^*L = F^* + G^*L_\alpha .$$

It follows that

$$T^{-1}F^*T + T^{-1}G^*[L + L_\alpha T - VL_\alpha] = F^*.$$

Hence, (T, V, L) belongs to the stabilizer of (F_s^*, G^*) if and only if $(T, V, L + L_\alpha T - VL_\alpha)$ belongs to the stabilizer of (F^*, G^*). By Corollary (5.4), $(T, V, L + L_\alpha T - VL_\alpha)$ belongs to the stabilizer of (F^*, G^*) if and only if

V = an $m \times m$ nonsingular matrix,

$T = \operatorname{diag}(V, V, \ldots, V)$,

$L + L_\alpha T - VL_\alpha = 0$.

However, if $T = \text{diag}(V, V, \ldots, V)$ then $L_\alpha T = VL_\alpha$. Thus, (T, V, L) belongs to the stabilizer of (F_s^*, G^*) if and only if

V = an $m \times m$ nonsingular matrix,

$T = \text{diag}(V, V, \ldots, V)$, $L = 0$. □

The above lemma shows that the stabilizer of the feedback group of the dual Kronecker canonical form (F^*, G^*) is the same as the stabilizer of the feedback group of (F_s^*, G^*).

Returning to the consideration of the Riccati group action on DS_1, we first have the following

(9.2) PROPOSITION. *Any quintuple* (F, G, Q, R, S) *in* D_s *can be transformed to a quintuple of the form* $(F_s^*, G^*, 0, I, \hat{S})$. *Further,* $(F_s^*, G^*, 0, I, S^1)$ *and* $(F_s^*, G^*, 0, I, S^2)$ *in* DS_1 *belong to the same orbit under the Riccati group if and only if there exists an* $m \times m$ *orthogonal matrix* V *such that*

$$\text{diag}(V', V', \ldots, V')S^1V = S^2.$$

PROOF. Choose a triple (T, V, L) such that

$$T^{-1}FT + T^{-1}GL = F_s^*, \quad T^{-1}GV = G^*.$$

Let Q_1, R_1, S_1 be such that

$$\psi((F, G, Q, R, S), (T, V, L, 0)) = (F_s^*, G^*, Q_1, R_1, S_1).$$

Since R_1 is positive definite, there exists an $m \times m$ nonsingular matrix V_1 such that $V_1'R_1V_1 = I$. With

$$T_1 = \text{diag}(V_1, V_1, \ldots, V_1),$$

we have

$$\psi((F_s^*, G^*, Q_1, R_1, S_1), (T_1, V_1, 0, 0)) = (F_s^*, G^*, Q_2, I, S_2),$$

for some Q_2 and S_2. Since the eigenvalues of F_s^* are in the left half plane, it is well known in the theory of Liapunov equation that there exists an $n \times n$ symmetric matrix M such that

$$(F_s^*)'M + MF_s^* = -Q_2.$$

Then

$$\psi((F_s^*, G^*, Q_2, I, S_2), (I, I, 0, M)) = (F_s^*, G^*, 0, I, S_2 - MG).$$

Combining all the transformations above, there exists $(\hat{T}, \hat{V}, \hat{L}, \hat{M})$ in the Riccati group such that

$$\psi((F, G, Q, R, S), (\hat{T}, \hat{V}, \hat{L}, \hat{M})) = (F_s^*, G^*, 0, I, \hat{S}),$$

where $\hat{S} := S_2 - MG$.

This proves the first part of the proposition.

Now, $(F_s^*, G^*, 0, I, S^1)$ and $(F_s^*, G^*, 0, I, S^2)$ belong to the same orbit under Γ if and only if there exists (T, V, L, M) in Γ such that

$$T^{-1}F_s^*T + T^{-1}G^*L = F_s^*, \tag{9.3}$$

$$T^{-1}G^*V = G^*, \tag{9.4}$$

$$T'((F_s^*)'M + MF_s^*)T + T'(S^1 + MG)L + L'(S^1 + MG)'T + L'L = 0, \tag{9.5}$$

$$(T'(S^1 + MG) + L')V = S^2, \tag{9.6}$$

$$V'V = I. \tag{9.7}$$

Equations (9.3) and (9.4) are equivalent to $T = \mathrm{diag}(V, V, \ldots, V)$, $L = 0$, and V any $m \times m$ nonsingular matrix. However, now (9.5) becomes

$$(F_s^*)'M + MF_s^* = 0.$$

Again, since the eigenvalues of F_s^* are in the open left half plane, it follows from the theory of Liapunov equation that $M = 0$. Finally, $V'V = I$ is equivalent to V being orthogonal. Thus, Equations (9.3) through (9.7) are equivalent to

$$T'S^1V = S^2$$

for some orthogonal matrix V and $T = \text{diag}(V, V, \ldots, V)$. Hence, $(F_s^*, G^*, 0, I, S^1)$ and $(F_s^*, G^*, 0, I, S^2)$ belong to the same orbit under Γ if and only if there exists an orthogonal matrix V such that

$$\text{diag}(V', V', \ldots, V')S^1V = S^2. \quad \square$$

We see that the problem of Riccati group action on DS_1 essentially reduces to the action of the orthogonal group on a space of matrices. We can now apply the results of Section 7 to get a complete set of invariants and canonical forms for the action of Γ on DS_1. Given any $n \times m$ matrix S, we partition it as

$$S' = [S_1', S_2', \ldots, S_r'].$$

Let (S) denote the r-tuple of $m \times m$ matrices $(S_1, S_2, \ldots, S_r)$. We can now state the following

(9.8) THEOREM. *Let* (F, G, Q, R, S) *be in* DS_1. *Then there exists an* $n \times m$ *matrix* $\hat{S}$ *such that* $(F_s^*, G^*, 0, I, \hat{S})$ *is in the same orbit as* (F, G, Q, R, S) *under the action of the Riccati group. Let* $TRS[(F, G, Q, R, S)]$ *denote* $\text{Tr}(\hat{S})$. *Then*

$$TRS: DS_1 \to \underline{\underline{R}}^{\eta}: (F, G, Q, R, S) \mapsto TR[(F, G, Q, R, S)]$$

is a well defined function (for some η). *Further,* d_1 *and* d_2 *in* DS_1 *belong to the same orbit under the Riccati group if and only if* $TRS[d_1] = TRS[d_2]$.

PROOF. The theorem follows immediately from Proposition (9.2) and (7.1). □

We note that using the techniques outlined at the end of Section 7 we can obtain canonical forms for S under the orthogonal group on certain dense open subsets.

The idea of using Liapunov equation generalizes to finite time problems as well. We state the result in

(9.9) THEOREM. *Let* $\hat{d} = (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ *be in* D *and let* $\chi(s)$ *be an arbitrary Hurwitz polynomial of degree* n. *Then there exists* $d = (F, G, 0, I, S, P_o)$ *in the orbit of* $\hat{d}$ *under* Γ *such that the characteristic polynomial of* F *is* χ.

PROOF. Let V be an $m \times m$ nonsingular matrix such that $V'\hat{R}V = I$. By well known results on state-feedback, there exists an $m \times n$ matrix L such that

$$\det (sI - F - GL) = \chi(s).$$

(See ROSENBROCK [1970] for a proof of this.) It now follows from the definition of φ that

$$\varphi(\hat{d}, (I, V, L, 0)) = (F, G, \tilde{Q}, I, \tilde{S}, \tilde{P}_o) =: \tilde{d}$$

for $\tilde{Q}, \tilde{S}, \tilde{P}_o$. It follows from the theory of Liapunov equation that there exists an $n \times n$ symmetric matrix M such that

$$\varphi(\tilde{d}, (I, I, 0, M)) = (F, G, 0, I, S, P_o)$$

for some S and P_o. □

For the case of equal Kronecker indices, we can obtain canonical forms and invariants following the techniques developed in this chapter. In general, however, this technique does not lead to canonical forms. Theorem (9.9) will turn out to be useful in Chapter VI.

CHAPTER V. ACTION OF THE RICCATI GROUP--II

In this chapter, we shall develop an alternative approach to obtaining invariants and canonical forms for the action of the Riccati group. This approach is based on the associated algebraic Riccati equation. Under the assumption that this algebraic Riccati equation has a real symmetric solution, we derive invariants and canonical forms. We get closed form results for invariants and canonical forms only for the cases of equal and strictly unequal Kronecker indices. The difficulties in the general case arise due to the fact that the stabilizer in the feedback group plays a very critical role - just as in Chapter IV. This stabilizer subgroup, in general, appears to be very complicated making it difficult to formulate any general results on invariants. Some of the results of this chapter have direct system theoretic applications which are discussed in the next chapter.

10. The Algebraic Riccati Equation Approach

We start by defining certain subsets of the optimization data spaces D and DS. For any quintuple (F, G, Q, R, S), recall that the algebraic Riccati equation (ARE) is

$$F'P + PF + Q - (PG + S)R^{-1}(PG + S)' = 0. \tag{10.1}$$

Now define

$$D_R := \{(F, G, Q, R, S, P_o) \text{ in } D: \text{ the ARE (10.1) associated with } (F, G, Q, R, S) \text{ has a real symmetric solution}\}.$$

In an analogous way, we have

$$DS_R := \{(F, G, Q, R, S) \text{ in } DS: \text{ the ARE (10.1) has a real symmetric solution}\}.$$

The key result in the subsequent development is the following

(10.2) PROPOSITION. Let $d := (F, G, Q, R, S, P_o)$ be in D. Then d belongs to D_R if and only if there exists $\gamma = (T, V, L, M)$ in the Riccati group Γ such that

$$\varphi(d, \gamma) = (\hat{F}, \hat{G}, 0, \hat{R}, 0, \hat{P}_o)$$

for some $\hat{F}$, $\hat{G}$, $\hat{R}$, and $\hat{P}_o$ such that the eigenvalues of $\hat{F}$ are in the closed left half plane.

PROOF. If d belongs to D_R then the ARE (10.1) has a real symmetric solution. Furthermore, it is known that in this case there exists a unique solution M such that with

$$L := - R^{-1}(MG + S)',$$

$(F + GL)$ has eigenvalues in the closed left half plane. (See WILLEMS [1971, Theorem 5].) Then

$$\hat{Q} = Q + F'M + MF + (S + MG)L + L'(S + MG)' + L'RL,$$
$$= Q + F'M + MF - (MG + S)R^{-1}(MG + S)' = 0$$

and also

$$\hat{S} = (S + MG) + L'R = 0.$$

Therefore, with $T = I$, $V = I$, we have

$$\varphi(d,(T, V, L, M)) = (\hat{F}, \hat{G}, 0, \hat{R}, 0, \hat{P}_o),$$

where

$$\hat{F} = F + GL = F - GR^{-1}(MG + S)',$$
$$\hat{R} = R, \quad \hat{G} = G, \quad \hat{P}_o = P_o - M.$$

Finally, M was chosen such that the roots of $\chi_{\hat{F}}$ are in the closed left half plane.

Conversely, suppose there exists $\gamma = (T, V, L, M)$ such that

$\varphi(d, \gamma) = (\hat{F}, \hat{G}, 0, \hat{R}, 0, \hat{P}_o)$ for some $\hat{F}, \hat{G}, \hat{R}, \hat{P}_o$. Then by the definition of φ, we have

$$(10.3)\quad \hat{Q} = T'(Q + F'M + MF)T + T'(S + MG)L + L'(S + MG)'T + L'RL = 0,$$

and

$$(10.4)\quad \hat{S} = (T'(S + MG) + L'R)V = 0.$$

Since V is nonsingular, (10.4) implies

$$(10.5)\quad L' = -T'(S + MG)R^{-1}.$$

Substituting (10.5) into (10.3), we have

$$\begin{aligned} 0 &= T'(Q + F'M + MF)T - T'(S + MG)R^{-1}(S + MG)'T - \\ &\quad - T'(S + MG)R^{-1}(S + MG)'T + T'(S + MG)R^{-1}(S + MG)'T, \\ &= T'(Q + F'M + MF - (S + MG)R^{-1}(S + MG)')T. \end{aligned}$$

As T is nonsingular, it follows that M is a solution to the ARE (10.1). Hence, d belongs to D_R. □

The above proposition may be regarded as the Riccati transformation group interpretation of the condition that the associated algebraic Riccati equation be solvable. We shall consider the consequences of this result on spectral theory of linear-quadratic optimal control problems as developed recently by JONCKHEERE and SILVERMAN [1978].

Thus, for each d in D_R, there exists a sixtuple $(F, G, 0, R, 0, P_o)$ in the orbit of d such that the eigenvalues of F are in the closed left half plane. Therefore for the analysis of the action of the Riccati group on D_R, we may restrict our attention to elements of the form $(F, G, 0, R, 0, P_o)$ such that the roots of χ_F are in the closed left half plane. The conditions under which any two such sixtuples belong to the same orbit under Γ are given in the following

(10.6) PROPOSITION. Let $d_1 = (F_1, G_1, 0, R_1, 0, P_{o1})$ and

$d_2 = (F_2, G_2, 0, R_2, 0, P_{o2})$ *be in* D_R *be such that the eigenvalues of* F_1 *and* F_2 *are in the closed left hand plane. Then* d_1 *and* d_2 *belong to the same orbit under* Γ *if and only if there exists an* $n \times n$ *nonsingular matrix* T *and an* $m \times m$ *nonsingular matrix* V *such that*

$$\begin{aligned} &T^{-1}F_1T = F_2, \quad T^{-1}G_1V = G_2, \\ &T'P_{o1}T = P_{o2}, \quad V'R_1V = R_2. \end{aligned} \tag{10.7}$$

PROOF. Suppose there exist nonsingular matrices T, V such that (10.7) holds. It then follows from the definition of φ that

$$\varphi((F_1, G_1, 0, R_1, 0, P_{o1}), (T, V, 0, 0)) = (F_2, G_2, 0, R_2, 0, P_{o2}).$$

Conversely, suppose there exists $\gamma = (T, V, L, M)$ in Γ such that

$$\varphi(d_1, \gamma) = d_2.$$

Then we have $V'R_1V = R_2$. Furthermore,

$$T'(F_1'M + MF_1)T + T'MG_1L + L'G_1'MT + L'R_1L = 0, \quad (T'MG_1 + L'R_1)V = 0.$$

As V is nonsingular, it follows that

$$L' = - T'MG_1R_1^{-1}.$$

Consequently

$$T'(F_1'M + MF_1 - MG_1R^{-1}G_2'M)T = 0.$$

As T is nonsingular, M satisfies

$$F_1'M + MF_1 - MG_1R_1^{-1}G_1'M = 0.$$

We also have

$$F_2 = T^{-1}F_1T + T^{-1}G_1L = T^{-1}(F_1 - G_1R_1^{-1}G_1'M)T.$$

As the eigenvalues of F_2 are the left half plane, it follows that the eigenvalues of $F_1 - G_1R_1^{-1}G_1'M$ are also in the closed left half plane. Thus, M is a real symmetric solution to the ARE (10.1) associated with $(F_1, G_1, 0, R_1, 0)$ such that the eigenvalues of $F_1 - G_1R_1^{-1}G_1'M$ are in the closed left half plane. Since the eigenvalues of F_1 are also in the closed left half plane, the $n \times n$ matrix consisting of zeroes is also one such solution. However, by the uniqueness of such a solution (see WILLEMS [1971, Theorem 5]), we must have $M = 0$. Consequently, $L = 0$. Thus,

$$T^{-1}F_1T = F_2, \quad T^{-1}G_1V = G_2, \quad T'P_{o1}T = P_{o2}, \quad V'R_1V = R_2. \quad \square$$

We thus see that every element of D_R can be brought into the form $(F, G, 0, R, 0, P_o)$ by a suitable element of the Riccati group. Further, such a form is unique up to state space transformations $(GL(n))$ and the input space transformations $(GL(m))$. We will now show that these actions are intimately related with the stabilizer of the feedback group. We state the result in the following

(10.8) PROPOSITION. *Let (F, G) be a reachable pair. Then there exist nonsingular matrices T, V such that*

$$T^{-1}FT = F^* - G^*L, \quad T^{-1}GV = G^*,$$

for some $m \times n$ matrix L. Furthermore, for any two $m \times n$ matrices L_1, L_2 there exist nonsingular matrices T^*, V^* such that

$$(10.9) \quad (T^*)^{-1}(F^* - G^*L_1)T^* = F^* - G^*L_2, \quad (T^*)^{-1}G^*V^* = G^*,$$

if and only if there exists L^* such that (T^*, V^*, L^*) belongs to the stabilizer in the feedback group of (F^*, G^*) and

$$L_2 = (V^*)^{-1}(L_1T^* + L^*).$$

PROOF. It follows from the results of Chapter II that there exists (T, V, L_3) in the feedback group such that

$$T^{-1}FT + T^{-1}GL_3 = F^*, \quad T^{-1}GV = G^*.$$

Defining $L := V^{-1}L_3$, we have

$$T^{-1}FT = F^* - G^*L, \quad T^{-1}GV = G^*.$$

Equations (5.9) hold if and only if

$$(T^*)^{-1}F^*T^* + (T^*)^{-1}G^*(V^*L_2 - L_1T^*) = F^*,$$
$$(T^*)^{-1}G^*V^* = G^*. \tag{10.10}$$

Now, (10.10) holds if and only if $(T^*, V^*, V^*L_2 - L_1T^*)$ belongs to the stabilizer in the feedback group of (F^*, G^*). Equivalently, (10.9) holds if and only if there exists (T^*, V^*, L^*) in the stabilizer in the feedback group such that

$$V^*L_2 - L_1T^* = L^*,$$

or

$$L_2 = (V^*)^{-1}(L_1T^* + L^*). \quad \square$$

Let us now return to the consideration of the action of the Riccati group Γ on D_R. Using the above proposition, we now have the following

(10.11) THEOREM. Let $\hat{d} = (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ be in D_R. Then there exist matrices L, R, P_o such that the eigenvalues of $F_L := F^* - G^*L$ are in the closed left half plane and there exists $\hat{\gamma}$ in Γ such that

$$\varphi(\hat{d}, \hat{\gamma}) = (F_L, G^*, 0, R, 0, P_o).$$

Further, any $d_1 = (F_1^* - G_1^*L_1, G_1^*, 0, R_1, 0, P_{o1})$ and $d_2 = (F_2^* - G_2^*L_2, G_2^*, 0, R_2, 0, P_{o2})$ in D_R, such that the eigenvalues of $F_1^* - G_1^*L_1$ and $F_2^* - G_2^*L_2$ are in the closed left half plane, belong to the same orbit under Γ if and only if the ordered Kronecker indices of $(F_1^* - G_1^*L_1, G_1^*)$ and $(F_2^* - G_2^*L_2, G_2^*)$ are the same (hence,

$F_1^* = F_2^* =: F^*$, $G_1^* = G_2^* = G^*$), and there exists (T^*, V^*, L^*) in the stabilizer in the feedback group of (F^*, G^*) such that

$$(10.12)\quad L_2 = (V^*)^{-1}(L_1T^* + L^*),$$

$$(10.13)\quad R_2 = (V^*)'R_1V^*,$$

$$(10.14)\quad P_{o2} = (T^*)'P_{o1}T^*.$$

PROOF. By Proposition (10.2), there exist $\tilde{F}$, $\tilde{G}$, $\tilde{R}$, $\tilde{P}_o$ such that $\tilde{d} = (\tilde{F}, \tilde{G}, 0, \tilde{R}, 0, \tilde{P}_o)$ belongs to the orbit of $\hat{d}$ under Γ. Further, by Proposition (10.8) there exist nonsingular matrices T, V such that

$$T^{-1}\tilde{F}T = F^* - G^*L, \quad T^{-1}\tilde{G}V = G^*.$$

Hence,

$$\varphi(\tilde{d}, (T, V, 0, 0)) = (F_L, G^*, 0, R, 0, P_o),$$

where $R = V'\tilde{R}V$ and $P_o = T'\tilde{P}_oT$. This proves the first part of the theorem.

By Proposition (10.6) d_1 and d_2 belong to the same orbit under Γ if and only if there exist nonsingular matrices T, V such that

$$(10.15)\quad \begin{aligned} &T^{-1}(F_1^* - G_1^*L_1)T = F_2^* - G_2^*L_2, \quad T^{-1}G_1^*V = G_2^*, \\ &T'P_{o1}T = P_{o2}, \quad V'R_1V = R_2. \end{aligned}$$

Hence the ordered Kronecker indices of $(F_1^* - G_1^*L_1, G_1^*)$ and $(F_2^* - G_2^*L_2, G_2^*)$ are the same. Hence $F_1^* = F_2^* = F^*$, $G_1^* = G_2^* = G^*$. Further, by Proposition (10.8) Equation (10.15) holds if and only if there exists (T^*, V^*, L^*) in the stabilizer in the feedback group of (F^*, G^*) such that

$$L_2 = (V^*)^{-1}(L_1T^* + L^*),$$

$$R_2 = (V^*)'R_1V^*,$$

$$P_{o2} = (T^*)'P_{o1}T^*. \quad \square$$

The above theorem shows that the problem of Riccati group action on D_R can be reduced to the action of the stabilizer of the feedback group on matrices L, R, P_o. As discussed in Chapter II, the stabilizer depends only on the Kronecker indices. This leads us to study the problem for various choices of Kronecker indices. Again, the case of equal Kronecker indices is the simplest to analyze as in Sections 6 and 8. Following the techniques developed in Section 8, we can also treat the case of strictly unequal Kronecker indices. In fact, the results here closely parallel the results of Sections 6 and 8.

Before analyzing the special cases, we have the following

(10.16) COROLLARY. Let $\hat{d} = (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ be in D_R. Then there exist matrices L, P_o such that the eigenvalues of $F_L := F^* - G^*L$ are in the closed left half plane and $d := (F_L, G^*, 0, I, 0, P_o)$ belongs to the orbit of $\hat{d}$.

PROOF. Note that for any set of ordered Kronecker indices, and any lower triangular matrix V^*, there exists T^* such that $(T^*, V^*, 0)$ belongs to the stabilizer of (F^*, G^*). By Theorem (10.11), we may choose $\hat{Q} = \hat{S} = 0$ and $\hat{F} = F^* - G^*L_1$, $\hat{G} = G^*$ for an appropriate L_1 such that the eigenvalues of $\hat{F}$ are in the left half plane. Now choose a lower triangular matrix V^* such that $V^{*\prime}\hat{R}V^* = I$. Then

$$\varphi((\hat{F}, G^*, 0, R, 0, \hat{P}_o), \ (T^*, V^*, 0, 0)) =$$

$$= (F_L, G^*, 0, I, 0, P_o)$$

for some P_o and

$$L = (V^*)^{-1} L_1 T^*.$$

Noting that $(T^*)^{-1} \hat{F} T^* = F_L$, the result follows. □

We start with the case of equal Kronecker indices. Let

$$D_{1R} := \{(F, G, Q, R, S, P_o) \text{ in } D_R: \text{ the Kronecker indices } (F, G) \text{ are equal}\}.$$

Given any $m \times n$ matrix L and $n \times n$ matrix P_o, let L and P_o be partitioned as

$$L = [L_1 \quad L_2 \quad \cdot \quad \cdot \quad \cdot \quad L_\ell],$$

$$P_o = \begin{bmatrix} P_{o11} & \cdot & \cdot & \cdot & P_{o1\ell} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ P_{o\ell 1} & & & & P_{o\ell\ell} \end{bmatrix},$$

where $\ell := n/m$. Now, by Theorem (10.11) for any $d_1 = (F_1, G_1, Q_1, R_1, S_1, P_{o1})$ in D_{1R}, there exist matrices L_2, R_2, P_{o2} such that

$$d_2 = (F^* - G^* L_2, G^*, 0, R_2, 0, P_{o2})$$

belongs to the orbit of d_1 under Γ. Let V be an $m \times m$ nonsingular matrix such that $V' R_2 V = I$. Let

$$T := \text{diag } (V, V, \ldots, V).$$

Then it is easy to see that

$$\varphi(d_2, (T, V, 0, 0)) = (F^* - G^* L, G^*, 0, I, 0, P_o),$$

where $L := V^{-1} L_2 T$ and $P_o := T' P_{o2} T$. We now define

$$TR_R[d_1] := Tr(P_o, L),$$

where (P_o, L) denote the collection of matrices P_{oij} and L_k. (We are following the notation developed in Chapter IV, Sections 7 and 8.) We can now state the main

(10.17) THEOREM. Let

$$TR_R: D_{1R} \to \underline{\underline{R}}^N: d \to TR_R[d].$$

Then TR_R is a well defined function. Further, for any d_1 and d_2 in D_{1R}, $TR_R[d_1] = TR_R[d_2]$ if and only if d_1 and d_2 belong to the same orbit under Γ.

PROOF. By the argument just before the theorem, there exists an element of the form $(F^* - G^*L, G^*, 0, I, 0, P_o)$ in every orbit of D_{1R} under Γ. Further, $d_1 = (F^* - G^*L_1, G^*, 0, I, 0, P_{o1})$ and $d_2 = (F^* - G^*L_2, G^*, 0, I, 0, P_{o2})$ in D_{1R} belong to the same orbit under Γ if and only if there exists (T^*, V^*, L^*) in the stabilizer of the feedback group of (F^*, G^*) such that

(10.18) $(V^*)'V^* = I$

and (10.12) and (10.14) hold. Since the Kronecker indices of (F^*, G^*) are equal, it follows from Corollary (5.4) and (10.18) that V^* is an $m \times m$ orthogonal matrix, $L^* = 0$, and

$$T^* = \text{diag}\,(V^*, V^*, \ldots, V^*).$$

We, thus, see that d_1 and d_2 belong to the same orbit under Γ if and only if

$$L_2 = (V^*)'L_1T^*,$$

$$P_{o2} = (T^*)'P_{o1}T^*,$$

where V^* is any $m \times m$ orthogonal matrix and $T^* = \text{diag}(V^*, V^*, \ldots, V^*)$. It now follows from Theorem (7.1) that TR_R is a well defined function and $TR_R[d_1] = TR_R[d_2]$ if and only if d_1 and d_2 belong to the same orbit under Γ. □

This theorem is very similar to Theorem (8.1) of the previous chapter. Essentially, the problem is the same: analyze the action of the orthogonal group on a space of matrices.

The case of strictly unequal Kronecker indices can be treated using the techniques developed in Section 8. Intuitively speaking, L plays the role of S' and P_o plays the role of Q. Here we state the results without proofs since the proofs follow the pattern of Section 8. We start by defining

$$D_{2R} := \{(F, G, Q, R, S, P_o) \text{ in } D_R: \text{ the Kronecker indices of } (F, G) \text{ are } \kappa_1 > \kappa_2 > \ldots \kappa_m \geq 1\}.$$

We now have

(10.19) THEOREM. Let $\hat{d} = (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ be any element of D_{2R}. Then there exist matrices P_o and L with L' in SC such that the eigenvalues of $F^* - G^*L$ are in the closed left half plane and $d = (F^* - G^*L, G^*, 0, I, 0, P_o)$ belongs to the orbit of $\hat{d}$ under the Riccati group. Further, the assignment

$$\ell p: D_{2R} \to QS: \hat{d} \mapsto [P_o, L']$$

is a well defined function. Finally, for any d_1 and d_2 in D_{2R}, $\ell p(d_1) = \ell p(d_2)$ if and only if d_1 and d_2 belong to the same orbit under the Riccati group.

Here we are following the notation of Section 8. Theorem (10.19) should be compared with Theorem (8.5). Roughly speaking, in the approach of the previous chapter, all the invariants occur in S and Q, whereas in the approach of this section the invariants occur in L and P_o.

Theorems (8.5) and (10.19) give a complete set of invariants in terms of finite collection of matrices.

We conclude this section with the case of infinite time-zero terminal state problems. We start with

(10.20) THEOREM. *Let* $d_1 = (F_1, G_1, Q_1, R_1, S_1)$ *be in* DS_R. *Then there exists an* $m \times n$ *matrix* L *and an* $m \times m$ *matrix* R *such that the eigenvalues of* $(F^* + G^*L)$ *are in the closed left half plane and* $d = (F^* + G^*L, G^*, 0, R, 0)$ *belongs to the same orbit as* d_1. *Further,* $d_1 = (F_1^* - G_1^*L_1, G_1^*, 0, R_1, 0)$ *and* $d_2 = (F_2^* - G_2^*L_2, G_2^*, 0, R_2, 0)$ *in* DS_R *such that the eigenvalues of* $(F_1^* - G^*L_1)$ *and* $(F_2^* - G_2^*L_2)$ *are in the closed left half plane, belong to the same orbit under* Γ *if and only if the ordered Kronecker indices of* (F_1^*, G_1^*) *and* (F_2^*, G_2^*) *are the same, (hence,* $F_1^* = F_2^* =: F^*$, $G_1^* = G_2^* =: G^*$), *and there exists* (T^*, V^*, L^*) *in the stabilizer of the feedback group of* (F^*, G^*) *such that*

$$(V^*)'R_1V^* = R_2.$$

$$L_2 = (V^*)^{-1}(L_1T^* + L^*).$$

PROOF. Recall that the Riccati group Γ acts on DS the same way as it does on D, except that the last term is omitted. With this observation, the proof follows from the proof of Theorem (10.11). □

Again, one can analyze the special cases of equal and strictly unequal Kronecker indices. The results are very similar to those obtained in Chapter IV, Section 9. We omit the details.

CHAPTER VI. SYSTEM THEORETIC APPLICATIONS

In this chapter, we will examine some of the system theoretic applications and interpretations of the Riccati transformation group and the techniques developed in the last two chapters. The main application of our results is to the spectral theory of linear-quadratic optimization problems, developed recently by JONCKHEERE and SILVERMAN [1978]. We will show that their results can be considerably extended and simplified using the Riccati transformation group. In particular, we will show by using the Riccati group that the boundedness of infima in the infinite time free terminal state case is equivalent to the positivity of an $n \times n$ symmetric matrix. We will show that the Hilbert space operators which are the central objects of study by JONCKHEERE and SILVERMAN [1978] are not invariant under the Riccati transformation group. We will examine the behavior of Popov's frequency function under the Riccati transformation group. We will also show that positive real transfer functions and the Riccati transformation group are closely related.

11. Spectral Theory of Linear Quadratic Optimal Control

A spectral theory of linear quadratic optimal control problems has recently been developed by JONCKHEERE and SILVERMAN [1978] in which they relate the problems of boundedness of the infima to the spectra of certain Hilbert space operators. Here we will examine these problems in the light of Riccati transformation group. We will analyze continuous-time versions of the discrete-time problems investigated by JONCKHEERE and SILVERMAN [1978]. Our techniques and results apply equally well to the discrete-time problems.

We start by precisely defining the dynamic optimization problems that we will consider. Following WILLEMS [1971], we define the infinite time free terminal state linear quadratic optimization problem associated with $d = (F, G, Q, R, S, P_0)$ in D as follows. Infimize the functional

(11.1) $$J_f(u(\cdot)) = \int_{-\infty}^{0}[x'(t)Qx(t) + 2x'(t)Su(t) + u'(t)Ru(t)]dt + x'(0)P_0x(0),$$

subject to the linear system

$$dx(t)/dt = Fx(t) + Gu(t),$$

$$\lim_{t\to-\infty} x(t) = 0,$$

over all $u(\cdot)$ in

$$L_2^m(-\infty, 0] := \{f: (-\infty, 0] \to \underline{R}^m: \|f(t)\| \text{ is square integrable on } (-\infty, 0]\}.$$

Here the infimum is to be taken as the infimum of limits. It turns out that the boundedness of the infima of infinite time free terminal state problems is of critical importance in the existence of solutions to the Riccati differential equations. We state the relevant results in the following

(11.2) THEOREM. Let $d = (F, G, Q, R, S, P_0)$ be in D. Then the following statements are equivalent.

(i) For the infinite time free terminal state problem associated with d,

$$\inf J_f(u(\cdot)) \geq 0,$$

where the infimum is taken over all u in $L_2^m(-\infty, 0]$.

(ii) The Riccati differential equation associated with d

(11.3) $$-dP(t)/dt = F'P(t) + P(t)F + Q - (P(t)G + S)R^{-1}(P(t)G + S)',$$
$$P(0) = P_0,$$

has a solution for all $t < 0$.

(iii) For all $t < 0$ and c in $\underline{R}^n$, the finite time linear quadratic optimization problem associated with d, t, c has a solution.

(iv) For all $t < 0$, the (zero initial state) finite time linear quadratic optimization problem associated with d, t, $c = 0$, has a solution.

(Sketch of the) PROOF. Equivalence between (ii) and (iii) follows essentially from the classical results of KALMAN [1960]. See ANDERSON and MOORE [1971, Section 2.3] for the details. Equivalence between (iii) and (iv) follows from the reachability of (F, G). The implication (iv) $\rightarrow$ (i) follows from a trivial modification of a result of WILLEMS [1971, Theorem 1].

We now prove (i) $\rightarrow$ (iv). Suppose that for some $t_o < 0$, the finite time linear quadratic optimization problem associated with d, t_o, and $c = 0$ does not have a solution. Then there exists a sequence $\{u_i\}_{i=1}^{\infty}$, u_i in $L_2^m[t_o, 0]$ such that

$$\lim_{i \to \infty} J(u_i(\cdot), t_o, 0) = -\infty.$$

Define

$$\hat{u}_i : (-\infty, 0] \rightarrow \underline{R}^m : \hat{u}_i(t) = \begin{cases} 0, & \text{for } t < t_o, \\ u_i(t), & \text{for } t \geq t_o. \end{cases}$$

Then $\hat{u}_i$ belongs to $L_2^m(-\infty, 0]$ for all i. Now with $x(t) = 0$ for $t < t_o$, we have

$$J_f(\hat{u}_i(\cdot)) = J(u(\cdot), t_o, 0).$$

Hence, $\inf J_f(u(\cdot)) = -\infty$. This contradicts (i). The result, thus, follows. □

This result is the main reason for the importance of study of infinite time free terminal state problems. WILLEMS [1971] and JONCKHEERE and SILVERMAN [1978] consider the restricted case of $P_o = 0$. Our results

of this section show that it is both natural and necessary to consider the general case of $P_o \neq 0$ in order to get a satisfactory theory. In fact, it turns out that P_o contains most of the critical information for the boundedness of infima.

The next theorem shows that the existence of solutions to the various dynamic optimization problems remains invariant under the action of the Riccati group.

(11.4) THEOREM. Let $d = (F, G, Q, R, S, P_o)$ be in D, $d_1 = (F, G, Q, R, S)$ be in DS, γ be in the Riccati group Γ, and $t_o < 0$. Then a solution to the infinite time free terminal state (respectively, the finite time, infinite time zero terminal state) linear-quadratic optimization problem associated with d (respectively, d and t_o, d_1) exists if and only if a solution to the same problem associated with $\varphi(d, \gamma)$ (respectively, $\varphi(d, \gamma)$ and t_o, $\psi(d_1, \gamma)$) exists.

PROOF. It follows from Theorem (11.2) that the infinite time free terminal state linear quadratic optimization problem associated with (F, G, Q, R, S, P_o) has a solution if and only if the associated Riccati differential equation (11.3) has a solution for all $t < 0$. However, if $P(t)$ is the solution to the Riccati differential equation associated with d then it is easy to check (by direct substitution) that $T'(P(t) - M)T$ is a solution to the Riccati differential equation associated with $\varphi(d, \gamma)$. Conversely, if $\hat{P}(t)$ is the solution to the Riccati differential equation associated with $\varphi(d, \gamma)$ then $(T')^{-1}\hat{P}(t)T^{-1} + M$ is the solution to the Riccati differential equation associated with d. It now follows from Theorem (11.2) that the infinite time free terminal state problem associated with d has a solution if and only if the same problem associated with $\varphi(d, \gamma)$ has a solution.

The proofs for the finite time problems and the infinite time zero terminal state problems are similar. □

The above result shows that in order to check the existence of solution to a dynamic optimization problem defined by a point d in D(or DS), we need to consider the problem at any one point on the

orbit of d. This simple observation allows us to extend and simplify the results of JONCKHEERE and SILVERMAN [1978] as we will demonstrate in the following results.

Let $\hat{d} = (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ be any point of D. Then, by Theorem (9.9) of Section 9, it follows that there exist matrices L, S, P_o such that $d = (F, G^*, 0, I, S, P_o)$ belongs to the orbit of $\hat{d}$ where $F := \hat{F}^* + \hat{G}^* L$ is asymptotically stable. Hence, in the light of Theorems (11.2) and (11.4), we need only consider points of the form $(F, G, 0, I, S, P_o)$ where F is asymptotically stable. (Note that we have dropped * from F, G as the actual canonical form of F, G does not play an important role in the following development.) Let us start with the infinite time free terminal state linear quadratic optimization problem associated with a six tuple in the form (F, G, 0, I, S, P_o), F asymptotically stable. Then the associated cost criterion is

$$(11.5)\quad J_f(u(\cdot)) = \int_{-\infty}^{0} [2x'(t)Su(t) + u'(t)Ru(t)]dt + x'(0)P_o x(0),$$

subject to the linear system

$$\dot{x}(t) = Fx(t) + Gu(t),$$

$$\lim_{t \to -\infty} x(t) = 0.$$

It is well known that the solution to the above equation is given by

$$x(t) = e^{F(t-t_o)}x(t_o) + \int_{t_o}^{t} e^{F(t-\tau)}Gu(\tau)d\tau.$$

As F is asymptotically stable, taking limit as $t_o \to -\infty$, we get

$$(11.6)\quad x(t) = \int_{-\infty}^{t} e^{F(t-\tau)}Gu(\tau)d\tau.$$

Substituting (11.6) into (11.5) we get

$$J_f(u(\cdot)) = \int_{-\infty}^{0} u'(t)(\rho_f u)(t)dt,$$

where

(11.7) ρ_f: $L_2^m(-\infty, 0] \to L_2^m(-\infty, 0]$,

(11.8) $\rho_f := \rho_1 + \rho_2 + \rho_3$,

(11.9) ρ_i: $L_2^m(-\infty, 0] \to L_2^m(-\infty, 0]$, $i = 1, 2, 3$;

(11.10) $(\rho_1 u)(t) := u(t)$,

(11.11) $(\rho_2 u)(t) := \int_{-\infty}^{0} w(t - \tau)u(\tau)d\tau$,

(11.12) $(\rho_3 u)(t) := G'e^{-F't}P_o\int_{-\infty}^{0} e^{-F\tau}Gu(\tau)d\tau$,

(11.13)
$$w(t) := \begin{cases} S'e^{Ft}G & t > 0 \\ G'e^{-F't}S & t < 0 \\ 0 & t = 0. \end{cases}$$

For any u, v in $L_2^m(-\infty, 0]$, let us define

$$\langle u, v \rangle := \int_{-\infty}^{0} u'(t)v(t)dt.$$

Then in this notation,

(11.14) $J_f(u) = \langle u, \rho_f u \rangle$

and ρ_f is the Hilbert space operator defined above. We note that $\rho_1 + \rho_2$ is a Wiener-Hopf type operator (with the "kernel" $R\delta(t) + w(t)$ whereas ρ_3 is a "Hankel like" operator. Thus, we have a decomposition of ρ_f as the sum of a Wiener-Hopf operator and a "Hankel like" operator. This decomposition corresponds to the decomposition obtained by JONCKHEERE and SILVERMAN [1978, Section 2-E] of the discrete-time analog of ρ_f as the sum of a Toeplitz operator and a Hankel like operator. Their proof involves some matrix manipulations and the system

theoretic significance of these decompositions remains unclear. In our case, it is clear from the definitions that $\rho_1 + \rho_2$ (or, the Wiener-Hopf operator) corresponds to the integral quadratic criteria whereas ρ_3 (or, the Hankel like operator) corresponds to final state weighing matrix P_o. In any case, the final state weighing matrix P_o is the natural tool to "account" for the so called "Hankel perturbation", (see JONCKHEERE and SILVERMAN [1978]).

We now have the following simple

(11.15) LEMMA. ρ is a bounded linear self adjoint operator.

PROOF. Clearly, ρ_f is a linear operator and ρ_1 is a bounded self adjoint. Since F is asymptotically stable, it follows from the definition of w that there exist $\alpha, \beta > 0$ such that

$$\|w(t)\| \leq \beta e^{-\alpha|t|}.$$

It then follows from GOHBERG and KREIN [1960, page 226] that ρ_2 is a bounded operator. As F is asymptotically stable, we can define

$$\mu := \|\int_{-\infty}^{o} e^{-Ft} GG' e^{-F't} dt\|.$$

It is easy to check that

$$\|\rho_3\| \leq \mu \|P_o\|.$$

Hence ρ_3 is also a bounded operator.

Since $w(t - \tau) = w'(\tau - t)$, it follows from YOSIDA [1964, page 197] that ρ_2 is self adjoint. Finally, for any u, v in $L_2^m(-\infty, 0]$,

$$\begin{aligned} \langle v, \rho_3 u \rangle &= \int_{-\infty}^{o} v'(t) G' e^{-F't} P_o \int_{-\infty}^{o} e^{-F\tau} Gu(\tau) d\tau dt \\ &= \int_{-\infty}^{o} u'(\tau) G' e^{-F'\tau} P_o \int_{-\infty}^{o} e^{-Ft} Gu(t) dt d\tau \\ &= \langle u, \rho_3 v \rangle. \end{aligned}$$

Hence ρ_3 is self adjoint. Thus, ρ_f is the sum of bounded linear self adjoint operators. Hence, ρ is a bounded linear self adjoint operator. □

We can now analyze the boundedness of the infima for infinite time free terminal state linear quadratic optimization problems in terms of the Hilbert space operator. In particular, we have

(11.16) THEOREM. Let $d = (F, G, Q, R, S, P_o)$ be in D such that F is asymptotically stable. Then the following statements are equivalent.

(i) The Riccati differential equation (11.3) associated with d has a solution for all $t < 0$.

(ii) The operator ρ_f associated with d is positive semidefinite.

PROOF. Note that $J_f(u) = \langle u, \rho_f u \rangle$ where ρ_f is a bounded linear self adjoint operator. Hence, by standard results on the spectral theory of self adjoint operators, $J_f \geq 0$ if and only if ρ_f is positive semidefinite. The conclusion follows from Theorem (11.2). □

This sets up the stage for a detailed investigation on the spectrum of ρ_f along the lines of JONCKHEERE and SILVERMAN [1978, 1980]. We will not pursue this avenue here as that is not our main purpose. In fact, a direct analysis of the spectrum of ρ_f seems to be difficult. JONCKHEERE and SILVERMAN [1978, 1980] overcome these difficulties by defining another Hilbert space operator whose spectrum differs from that of ρ_f only in zero. However, in the light of Theorem (11.4), it is not necessary to compute the spectrum of ρ_f associated with $(F, G, 0, I, S, P_o)$ itself. It may be easier to check the positivity of ρ_f at some other point on the orbit of d. As a matter of fact, this turns out to be the case.

Let $\hat{d} = (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ be in D. It follows directly from the results of WILLEMS [1971, Theorems 3, 4, 5] that it is necessary for the associated algebraic Riccati equation (ARE)

$$(11.17)\quad \hat{F}'\hat{P} + \hat{P}\hat{F} + \hat{Q} - (\hat{P}\hat{G} + \hat{S})\hat{R}^{-1}(\hat{P}\hat{G} + \hat{S})' = 0$$

to admit a real symmetric solution in order that the Riccati differential equation (11.3) associated with $\hat{d}$ have a solution for all $t < 0$. However, if the algebraic Riccati equation (11.17) has a real symmetric solution, then we are in the setup of Chapter V. In particular, Corollary (10.16) implies that there exist matrices L, P_o such that $(F, G, 0, I, 0, P_o)$ belongs to the orbit of $\hat{d}$ such that $F := F^* + G^*L$ is stable, (but may not be asymptotically stable). However, in order to use the operator ρ_f, we need to assume that F is asymptotically stable. We will first restrict our attention to those points $\hat{d}$ for which the resulting F turns out to be asymptotically stable. We will give operator theoretic interpretations of these considerations later in this section.

So, let $d = (F, G, 0, I, 0, P_o)$ be in D such that F is asymptotically stable. It follows directly from our previous results that the operator ρ_f in this case is

$$(11.18)\quad \rho_f = \rho_1 + \rho_3,$$

where ρ_1 and ρ_3 are as in (11.10) and (11.12). We will now show that ρ_f is positive semidefinite (respectively, definite), if and only if a certain $n \times n$ symmetric matrix is positive semidefinite (respectively, definite). The main result is based on the following simple

(11.19) LEMMA. *Let H_1 and H_2 be any two Hilbert spaces and let $\alpha: H_1 \to H_2$ and $\beta: H_2 \to H_1$ be any two bounded linear operators. Let $I_1: H_1 \to H_1$ and $I_2: H_2 \to H_2$ be the identity operators. Then, for any complex number $\lambda \neq 1$, λ belongs to the spectrum of $I_1 + \beta\alpha$ if and only if λ belongs to the spectrum of $I_2 + \alpha\beta$.*

PROOF. Let $\lambda \neq 1$ be a complex number in the spectrum of $I_1 + \beta\alpha$. Let $\delta := (\lambda - 1)^{-1}$. Now, suppose λ does not belong to the spectrum of $I_2 + \alpha\beta$. Then $(\lambda - 1)I_2 - \alpha\beta$ is an invertible operator. It follows that $I_2 - \delta\alpha\beta$ is also invertible. It is easy to check (by direct verification) that

$$[(\lambda - 1)I_1 - \beta\alpha][I_2 - \delta\beta(I_2 - \delta\alpha\beta)^{-1}\alpha]\delta = I_1.$$

But this implies that $(\lambda - 1)I_2 - \beta\alpha$ is invertible. This contradicts the fact λ is in the spectrum of $I_1 + \beta\alpha$. Hence, λ must belong to the spectrum of $I_2 + \alpha\beta$. The converse follows by a symmetrical argument. □

We now return to the analysis of the operator $\rho_1 + \rho_3$. Notice that the operator ρ_3 given in (11.12) can be written as

(11.20) $\rho_3 = \rho_4\rho_5$,

where

(11.21) $\rho_5: L_2^m(-\infty, 0] \to \underline{\underline{R}}^n: u \mapsto P_o\int_{-\infty}^{o} e^{-Ft}Gu(t)\,dt$,

(11.22) $\rho_4: \underline{\underline{R}}^n \to L_2^m(-\infty, 0]: x \mapsto \rho_4 x,\ (\rho_4 x)(t) := G'e^{-F't}x$.

We now apply the result of Lemma (11.19) to check the positivity of $\rho_1 + \rho_3$.

(11.23) THEOREM. Let $d = (F, G, 0, I, 0, P_o)$ be in D such that F is asymptotically stable. Define the matrix

(11.24) $\pi := \int_{-\infty}^{o} e^{-Ft}GG'e^{-F't}\,dt$.

Then, the operator $\rho_f (= \rho_1 + \rho_3)$ associated with d is positive semidefinite (respectively, definite) if and only if the eigenvalues of $I + P_o\pi$ are nonnegative (respectively, positive).

PROOF. Note that ρ_1 is the identity operator on $L_2^m(-\infty, 0]$ and $\rho_3 = \rho_4\rho_5$, where ρ_4, ρ_5 are given by (11.21), (11.22). As F is asymptotically stable, ρ_4, ρ_5 are bounded linear operators. We can now apply the result of Lemma (11.19). It follows that any (real) number $\lambda \neq 1$ belongs to the spectrum of ρ_f if and only if it belongs to the spectrum of

$$I + \rho_5\rho_4: \underline{\underline{R}}^n \to \underline{\underline{R}}^n.$$

However,

$$(\rho_5\rho_4)(x) = P_o\int_{-\infty}^{o} e^{-Ft}GG'e^{-F't}dt \quad x = P_o\pi x.$$

Consequently, the spectrum of ρ_f consists of the eigenvalues of the matrix $I + P_o\pi$ modulo 1. Hence ρ_f is positive semidefinite (respectively, definite) if and only if the eigenvalues of $(I + P_o\pi)$ are nonnegative (respectively, positive). □

Thus, we see that in order to check the positivity of the operator ρ_f associated with a six tuple $d = (F, G, O, I, O, P_o)$, it is necessary and sufficient to check the positivity of the eigenvalues of $I + P_o\pi$.

Even though P_o and π are symmetric matrices, $P_o\pi$ may not be a symmetric matrix. Therefore, a priori it is not obvious that $I + P_o\pi$ has real eigenvalues. Note that by reachability of (F, G), π is positive definite. Further, we may write

$$I + P_o\pi = (\pi^{-1} + P_o)\pi.$$

Thus, $I + P_o\pi$ can be written as a product of a symmetric matrix and a positive definite matrix. We can now use the classical results on regular pencils of quadratic forms. The reader is referred to GANTMACHER [1959, Chapter 10, Section 6, 7]. In particular, we have

(11.25) PROPOSITION. Let P_o, π be $n \times n$ real symmetric matrices and π be positive definite. Then, the eigenvalues of $(I + P_o\pi)$ are nonnegative (respectively, positive) if and only if the $n \times n$ real symmetric matrix $(\pi^{-1} + P_o)$ is positive semidefinite (definite).

PROOF. We can write

$$(I + P_o\pi) = (\pi^{-1} + P_o)\pi.$$

Now, it follows from GANTMACHER [1958, Chapter 10, Section 6, Theorem 8] that $(I + P_o\pi)$ has simple structure and real eigenvalues. Further,

it follows from GANTMACHER [1959, Chapter 10, Section 7, Theorem 10] that the eigenvalues of $(I + P_o\pi)$ are nonnegative (respectively, positive) if and only if $(\pi^{-1} + P_o)$ is positive semidefinite (definite). □

Note that $(\pi^{-1} + P_o)$ is a real symmetric matrix. We now summarize some of the main results of this section in the following

(11.26) THEOREM. Let $\hat{d} = (\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ be in D. Then the Riccati differential equation (11.3) associated with $\hat{d}$ has a solution for all $t < 0$ only if the associated algebraic Riccati equation has a real symmetric solution. Now, suppose that the algebraic Riccati equation associated with $\hat{d}$ has a real symmetric solution. Then there exist F, G, P_o such that $d = (F, G, 0, I, 0, P_o)$ belongs to the orbit of $\hat{d}$ and F is stable. In case F is asymptotically stable, then the Riccati differential equation associated with $\hat{d}$ has a solution for all $t < 0$ if and only if the $n \times n$ real symmetric matrix $(\pi^{-1} + P_o)$ is positive semidefinite, where

$$\pi := \int_{-\infty}^{0} e^{-Ft} GG' e^{-F't} dt.$$

Some remarks are in order here. The transformation from $(\hat{F}, \hat{G}, \hat{Q}, \hat{R}, \hat{S}, \hat{P}_o)$ to $(F, G, 0, I, 0, P_o)$ mainly requires a real symmetric solution to the algebraic Riccati equation and finding a square root of a positive definite matrix. Of course, there is no known finite procedure to find solutions to the algebraic Riccati equation. However, there are several efficient algorithms to actually compute the stabilizing solution to the algebraic Riccati equation. In any case, our results completely clarify and extend the investigation started by JONCKHEERE and SILVERMAN [1978]. We note that our condition that F be asymptotically stable is satisfied generically in D_R, the data space corresponding to those points of D for which the algebraic Riccati equation has a real symmetric solution.

Above results have a very simple control theoretic interpretation in terms of minimal energy control. In fact, we can derive a generalization of Theorem (11.26) using some classical results on minimal energy

control by KALMAN, HO, and NARENDRA [1963]. We now proceed to describe these considerations. Let $d = (F, G, 0, I, 0, P_o)$ be in D. Now consider the (zero initial state) finite time linear quadratic optimization problem associated with d, t_o, $c = 0$, (for any $t_o < 0$). Minimize

$$J(u(\cdot), 0, t_o) = \int_{t_o}^{o} u'(t)u(t)dt + x'(0)P_o x(0),$$

subject to the linear system

$$dx(t)/dt = Fx(t) + Gu(t), \quad x(t_o) = 0.$$

Clearly, the first term in the cost criterion corresponds to "control energy" whereas the second term corresponds to a weighing on the final state $x(0)$. Now, if $P_o = 0$, then the solution to this problem, obviously, is $u = 0$! If $P_o \neq 0$ then the above problem can be solved using the results on minimal energy control as developed by KALMAN, HO, and NARENDRA [1963]. Let us define, for each $t < 0$

$$\pi(t) = \int_t^o e^{-F\tau} GG' e^{-F'\tau} d\tau.$$

Since (F, G) is reachable, it follows that $\pi(t)$ is positive definite for all $t < 0$. Let x_o be an arbitrary but fixed vector in $\underline{\underline{R}}^n$. Consider the set $U(t, x_o)$ of input functions u in $L_2^m[t, 0]$ such that with $x(t_o) = 0$, the state at time zero $x(0) = x_o$, i. e., u transfers the system from the state 0 at t_o to the state x_o at time 0. Then Corollary 1 of Section 5 of KALMAN, HO, and NARENDRA [1963] implies that

$$\min \int_t^o u'(\tau)u(\tau)d\tau = x_o'\pi(t)^{-1}x_o.$$

We can now state the following

(11.27) THEOREM. Let (F, G) be a reachable pair. Then for any $t_2 \leq t_1$, we have that $\pi(t_2) \leq \pi(t_1)$, and

$$\lim_{t \to -\infty} \pi(t)^{-1} \quad (=: \hat{\pi})$$

exists. If F is asymptotically stable then $\hat{\pi} = \pi^{-1}$.

PROOF. Let $t_2 < t_1 \leq 0$. Let x_o be any arbitrary but fixed vector in $\underline{\underline{R}}^n$. For any u in $U(t_1, x_o)$ define

$$\hat{u}: L_2^m[t_2, 0] \to \underline{\underline{R}}^m$$

by $\hat{u}(t) = 0$ for $t_2 \leq t < t_1$ and $\hat{u}(t) = u(t)$ for $t_1 \leq t \leq 0$. It is evident that for each u in $U(t_1, x_o)$, $\hat{u}$ belongs to $U(t_2, x_o)$. It now follows that

$$\min_{U(t_2, x_o)} \int_{t_2}^{o} u'(\tau)u(\tau)d\tau \leq \min_{U(t_1, x_o)} \int_{t_1}^{o} u'(\tau)u(\tau)d\tau.$$

Now, from the result of KALMAN, HO, and NARENDRA [1963, Corollary 1, Section 5] quoted above, we have

$$x_o'\pi(t_2)^{-1}x_o \leq x_o'\pi(t_1)^{-1}x_o.$$

The above inequality holds for all x_o in $\underline{\underline{R}}^n$. Since $\pi(t)$ is positive definite for all t, it follows that

$$x_o'\pi(t)^{-1}x_o \geq 0.$$

Thus, $x_o'\pi(t)^{-1}x_o$ is a monotonically increasing function of t bounded below by 0. Therefore

$$\lim_{t \to -\infty} x_o'\pi(t)^{-1}x_o$$

exists for each x_o in $\underline{\underline{R}}^n$. Consequently,

$$\lim_{t \to -\infty} \pi(t)^{-1} \quad (=: \hat{\pi})$$

exists.

If F is asymptotically stable, then by the definition of π

$$\lim_{t\to-\infty} \pi(t) = \pi > 0.$$

By elementary results of analysis, we may conclude that

$$\hat{\pi} = \lim_{t\to-\infty} \pi(t)^{-1} = \pi^{-1}. \quad \square$$

We are now ready to state our main

(11.28) THEOREM. Let $\tilde{d} = (\tilde{F}, \tilde{G}, \tilde{Q}, \tilde{R}, \tilde{S}, \tilde{P}_o)$ be in D. Then, a solution to the Riccati differential equation exists for all $t < 0$ only if the algebraic Riccati equation associated with $\tilde{d}$ has a solution. Suppose that the algebraic Riccati equation associated with $\tilde{d}$ has a solution. Then there exist F, G, P_o such that $d = (F, G, 0, I, 0, P_o)$ belongs to the orbit of $\tilde{d}$. Furthermore, the Riccati differential equation associated with $\tilde{d}$ has a solution for all $t < 0$ if and only if $n \times n$ real symmetric matrix $(\hat{\pi} + P_o)$ is positive semi-definite.

PROOF. The first part of the theorem follows from Theorem (11.26). The existence of F, G, P_o such that $d = (F, G, 0, I, 0, P_o)$ belongs to the orbit of $\tilde{d}$ follows from Corollary (10.16) of Chapter V. It now follows from Theorem (11.4) that the Riccati differential equation associated with $\tilde{d}$ has a solution for all $t < 0$ if and only if the Riccati differential equation associated with d has a solution for all $t < 0$. Now, by Theorem (11.2), the Riccati differential equation associated with d has a solution for all $t < 0$ and only if for all $t < 0$, the (zero initial state) finite time linear quadratic optimization problem associated with d, t, $c = 0$ has a solution. The cost criteria in this case is

$$J(u(\cdot), t, 0) = \int_t^0 u'(\tau)u(\tau)d\tau + x'(0)P_o x(0).$$

It now follows from the result of KALMAN, HO, and NARENDA [1963, Section 5, Corollary 1] mentioned above that

$$\inf_{u\in L_2^m[t,0]} J(u(\cdot),\, t,\, 0) = \inf_{x\in R^n} x'(\pi(t)^{-1} + P_o)x.$$

Hence, this finite time optimization problem has a solution if and only if $(\pi(t)^{-1} + P_o)$ is positive semidefinite. Theorem (11.27) implies that for any $t_2 \le t_1 \le 0$,

$$\pi(t_2)^{-1} + P_o \le \pi(t_1)^{-1} + P_o.$$

Thus, again by Theorem (11.27), $(\pi(t)^{-1} + P_o)$ is positive semidefinite for all $t < 0$ if and only if $(\hat{\pi} + P_o)$ is positive semidefinite. Consequently, the Riccati differential equation associated with d has a solution if and only $(\hat{\pi} + P_o)$ is positive semidefinite. (In particular, if F is asymptotically stable then $(\hat{\pi} + P_o) = (\pi^{-1} + P_o)$.) $\square$

Note that Theorems (11.27) and (11.28) together imply the results of Theorem (11.26). Roughly speaking, we conclude that Riccati differential equation decomposes into two parts; one part consists of the algebraic Riccati equation (or equivalently, Wiener-Hopf operators), the other part consists of minimum energy control with final state weighing. The conditions for solvability of the algebraic Riccati equation are well understood. See, for example, WILLEMS [1971]. On the other hand, it is quite easy to check positive semidefiniteness of an $n \times n$ symmetric matrix. This can be done, for example, by computing all the principal minors.

(11.29) REMARK. Let $d = (F, G, Q, R, S, P_o)$ be in D. It was noted before that the Riccati differential equation (11.3) associated with d has a solution for all $t < 0$ only if the algebraic Riccati equation associated with d has a real symmetric solution. This result has a natural operator theoretic interpretation. Here we will describe the relevant results without proofs. The proofs may be constructed along the lines of JONCKHEERE and SILVERMAN [1978] using the standard techniques of functional analysis. It follows from Theorem (9.9) of Section 9 that each orbit in D contains a point in the form $d = (F, G, 0, I, S, P_o)$

where F is asymptotically stable. Then, the associated operator ρ_f is the sum of $(\rho_1 + \rho_2)$ and ρ_3 given by (11.10) - (11.12). $\rho_1 + \rho_2$ is a Wiener-Hopf operator. Using the asymptotic stability of F, it can be shown that ρ_3 is a completely continuous operator of the trace class. It can then be proved by techniques from perturbation theory that the spectrum of $(\rho_1 + \rho_2)$ is contained in the spectrum of ρ_f. Therefore, in order that ρ_f be positive semidefinite, it is necessary that $(\rho_1 + \rho_2)$ be positive semidefinite. Now by the theory of Wiener-Hopf operators, (for which see GOHBERG and KREIN [1960] and ROSENBLUM [1965]) the spectrum of $(\rho_1 + \rho_2)$ is given by the eigenvalues of the "frequency function"

$$\Phi(\omega) = I + S'(i\omega I - F)^{-1}G + G'(- i\omega I - F')^{-1}S ,$$

for all ω in $\underline{\underline{R}}$. Hence, $(\rho_1 + \rho_2)$ is positive semidefinite if and only if

$$\Phi(\omega) \geq 0 ,$$

for all ω in $\underline{\underline{R}}$. The nonnegativity of $\Phi(\omega)$ for all ω in $\underline{\underline{R}}$ is well known to be equivalent to the solvability of the algebraic Riccati equation associated with d. We, thus, see that in order that ρ_f be positive semidefinite it is necessary that the associated algebraic Riccati equation have a solution.

We close this section by examining the behavior of Popov's frequency domain function

$$\Phi_d(s) := R + S'(sI - F)^{-1}G + G'(- sI - F')^{-1}S +$$

$$+ G'(- sI - F')^{-1}Q(sI - F)^{-1}G$$

associated with $d = (F, G, Q, R, S)$ in DS under the action of the Riccati group. Let $\gamma = (T, V, L, M)$ be in Γ and let $\hat{d} := \psi(d, \gamma)$. Then it is a routine direct verification that

$$(11.30)\quad \Phi_{\hat{d}}(s) = V'(I - G'(-sI - F')^{-1}L')\Phi_d(s)(I - L(sI - F)^{-1}G)V.$$

In case, the algebraic Riccati equation associated with d has a real symmetric solution, then there exist matrices T, V, L, M such that $\hat{d} = (F, G, O, I, O)$, and

$$\Phi_{\hat{d}}(s) = I.$$

Equation (11.30) shows that this corresponds to a spectral factorization of $\Phi_{\hat{d}}(s)$ which is well known to be closely related to algebraic Riccati equation.

12. Other Applications

This section is mainly intended to show that there is a close relationship between the Riccati transformation group and positive real matrices. We show that the so called "positive real lemma" may be rederived from the Riccati transformation group point of view. We also obtain a complete parametrization of scalar positive real functions. Finally, we make some remarks on the fast algorithms for solving the Riccati differential equation due to CASTI [1974] and KAILATH [1973].

Let $Z(s)$ be an $m \times m$ matrix of real rational functions. Then $Z(s)$ is said to be positive real iff

1. all elements of $Z(s)$ are analytic in the region $Re(s) > 0$ of the complex plane, and
2. the matrix $Z^*(s) + Z(s)$ is positive semidefinite for all complex numbers s such that $Re(s) > 0$, where $Z^*(s)$ is the complex conjugate transpose of $Z(s)$.

The property of positive realness turns out to be of critical importance in passive network synthesis. (See NEWCOMB [1966] for the network theoretic aspects of positive real functions.) Using certain preliminary synthesis steps, (see ANDERSON and VONGPANITLERD [1973, Section 5.4],) it can be shown that the following assumptions on $Z(s)$ do not cause any considerable loss of generality.

1. $Z(\infty) := \lim_{s \to \infty} Z(s)$ exists. In other words, $Z(s)$ is a proper transfer function. Further, $Z(\infty) + Z'(\infty)$ is nonsingular.
2. $Z(s)$ is analytic is $\mathrm{Re}(s) \geq 0$.

Since $Z(s)$ is rational, using elementary realization theory, the above assumptions are easily seen to be equivalent to the existence of matrices F, G, H, J of appropriate sizes such that

$$Z(s) = J + H(sI - F)^{-1}G,$$

where (F, G, H) is canonical, $J + J'$ is nonsingular, and F is asymptomatically stable. In the light of these considerations, we may restate the problem as follows. Let F, G, H, J be $n \times n$, $n \times m$, $m \times n$, $m \times m$ real matrices respectively. Let (F, G, H) be canonical, F be asymptotically stable, and $J + J'$ be nonsingular. Find conditions on (F, G, H, J) such that

$$\Phi(s) := J + J' + H(sI - F)^{-1}G + G'(s^{*}I - F')^{-1}H' \geq 0, \tag{12.1}$$

for all s in the complex plane such that $\mathrm{Re}(s) \geq 0$. It is well known that in this case the positive semidefiniteness in the closed right half plane may be replaced by

$$\Phi(i\omega) \geq 0, \tag{12.2}$$

for all ω in $\underline{R}$. Let $R := J + J'$. Then a necessary condition for (12.1) to hold is that R be positive definite. However, notice that $\Phi(i\omega)$ is the Popov's frequency function associated with $(F, G, 0, R, H')$ in DS. Hence, by a result due to WILLEMS [1971, Theorem 5], we see that $Z(s)$ is positive real if and only if R is positive definite and the algebraic Riccati equation associated with $(F, G, 0, R, H')$ has a real symmetric solution. But now we are in the set up of Chapter V. In particular, the proof of Proposition (10.2) implies that the algebraic Riccati equation associated with $d = (F, G, 0, R, H')$ has a real symmetric solution if and only if there exists $\gamma = (I, V, L, M)$ such that

$$\psi(d, \gamma) = (\hat{F}, \hat{G}, 0, I, 0) \tag{12.3}$$

for some $(\hat{F}, \hat{G})$ such that the eigenvalues of $\hat{F}$ are in the left half plane. Rewriting (12.3) explicitly, we get

(12.4) $V'RV = V'(J + J')V = I,$

(12.5) $F'M + MF = L'RL,$

(12.6) $H' + MG = -L'R,$

(12.7) $\hat{F} = F + GL, \quad \hat{G} = GV.$

Equations (12.4) - (12.7) are easily seen to be equivalent to the so called "positive real lemma" equations due to YAKUBOVIC [1962], KALMAN [1963], and ANDERSON [1967]. This shows that the positive real lemma is closely related to the Riccati transformation group.

We now consider the special case of scalar positive real functions. In this case, J may be normalized to $1/2$. Thus,

(12.8) $Z(s) = 1/2 + h(sI - F)^{-1}g.$

Further, without loss of generality (F, g) may be taken to be in control canonical form:

(12.9) $$F := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ a_n & a_{n-1} & \cdots & & a_1 \end{bmatrix}, \quad g := \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}.$$

Then the characteristic polynomial of F is

(12.10) $\chi_F(s) := \det(sI - F) = s^n - a_1 s^{n-1} \ldots - a_n.$

We now consider the following problem. Let $a_1, a_2, \ldots, a_n$ be any fixed real numbers such that with (F, g) in the form (12.9), F is asymptotically stable. Then describe the set of all those $1 \times n$

matrices h such that $Z(s)$ given by (12.8) is positive real. Choose any $1 \times n$ vector ℓ such that the eigenvalues of $(F + g\ell)$ are in the (closed) left half plane. Let M be the unique solution to the Liapunov equation

(12.11) $F'M + MF = \ell'\ell$.

Existence and uniqueness of M follows from asymptotic stability of F. (In fact, $M = \int_0^\infty e^{F'T}\ell'\ell e^{FT}dt$.) Now define

(12.12) $h_\ell := -\ell - g'M$.

It is easy to verify that

$$Z(s) = 1/2 + h_\ell(sI - F)^{-1}g$$

is a positive real function. We can now state the following

(12.13) THEOREM. *Let $a_1, a_2, \ldots, a_n$ be any fixed real numbers, F, g be defined by (12.9), and F be asymptotically stable. Define the set*

$$C := \{h \text{ in } \underline{\underline{R}}^{1\times n}: Z(s) = 1/2 + h(sI - F)^{-1}g \text{ is positive real}\}.$$

Let Λ be the set

$$\Lambda := \{\ell \text{ in } \underline{\underline{R}}^{1\times n}: \text{the eigenvalues of } F + g\ell \text{ are in the (closed) left half plane}\}.$$

Then

$$\hat{h}: \Lambda \to C: \ell \mapsto h_\ell$$

is a bijection.

PROOF. From the arguments given above, it is clear that $\hat{h}$ is well defined. Suppose h belongs to C. Then there exist matrices ℓ, M such that

$$F'M + MF = \ell'\ell,$$

$$h = -\ell - g'M,$$

and the eigenvalues of $F + g\ell$ are in the closed left half plane. Hence, ℓ belongs to Λ and $h = h_\ell$. This proves that $\hat{h}$ is surjective.

Now suppose that there exist ℓ_1, ℓ_2 in Λ such that

$$\hat{h}(\ell_1) = \hat{h}(\ell_2) = h.$$

Then there exist symmetric matrices M_1, M_2 such that

$$\ell_i = -h - g'M_i, \quad i = 1, 2,$$

and

$$F'M_i + M_iF = \ell_i'\ell_i, \quad i = 1, 2.$$

Hence, M_1, M_2 are solutions to the algebraic Riccati equation

$$F'P + PF - (Pg + h')(Pg + h')' = 0,$$

such that the eigenvalues of $F + g\ell_1$ and $F + g\ell_2$ both are in the closed right half plane. However, by WILLEMS [1971, Theorem 5] such a solution is unique, and, hence, $\ell_1 = \ell_2$. This proves that $\hat{h}$ is injective. □

This theorem completely answers the question raised above. We note that the main step in obtaining a formula for $\hat{h}$ involves the solution of the (linear) Liapunov equation (12.11). Using Kronecker products we can obtain a closed form expression for M as a function

of ℓ. Combining this formula for M with (12.12), we can obtain a closed form expression for h_ℓ. Λ, as a subset of $\underline{\underline{R}}^n$, is the set of all ℓ such that the eigenvalues of $F + g\ell$ are in the closed left half plane. Using the classical stability theory and Hurwitz determinants, Λ can be described by a system of polynomial inequalities. (See, also, KALMAN [1979].) We note that there may exist ℓ in Λ such that (h_ℓ, F) is not observable. ($\ell = 0$ is one such choice.) Define

$$\Lambda_o := \{\ell \text{ in } \Lambda: (h_\ell, F) \text{ is observable}\}.$$

Then, ℓ belongs to Λ_o if and only if ℓ belongs to Λ and $(-\ell - g'M, F)$ is observable. Thus, Λ_o may be described as a subset of $\underline{\underline{R}}^n$ by some suitable polynomial inequalities.

We close this section with a remark on the fast algorithms for solving Riccati differential equation due to KAILATH [1973] and CASTI [1974]. Let $d = (F, G, Q, R, S, P_o)$ be in D. Consider the Riccati differential equation (11.3) associated with d. The main observation of KAILATH [1973] and CASTI [1974] is that if

$$r := \operatorname{rank} \dot{P}(0) = \operatorname{rank} (F'P_o + P_oF + Q - (P_oG + S)R^{-1}(P_oG + S)'),$$

then the Riccati differential equation may be replaced by a set of $n(r + m)$ nonlinear differential equations. In particular, if r is much smaller than n, then this leads to a considerable saving of computation. It turns out that this rank r is <u>invariant</u> under the action of the Riccati transformation group. This shows that one cannot achieve any further reduction of Riccati equation along the lines of KAILATH [1973] and CASTI [1974] using Riccati transformation group. However, a completely satisfactory explanation of this reduction still remains elusive.

REFERENCES

B. D. O. ANDERSON

[1967] "A system theory criterion for positive real matrices", SIAM J. Control, 5: 171-182.

B. D. O. ANDERSON and J. B. MOORE

[1971] Linear Optimal Control, Prentice-Hall, Englewood Cliffs, NJ.

B. D. O. ANDERSON and S. VONGPANITLERD

[1973] Network Analysis and Synthesis, Prentice-Hall, Englewood Cliffs, NJ.

P. BERNHARD and G. COHEN

[1973] "Etude d'une fonction fréquentielle intervenant dans un problème de commande optimale avec une application à la réduction de la taille de ce problème", Revue RAIRO, July 1973, pages 63-85.

R. W. BROCKETT

[1970] Finite Dimensional Linear Systems, Wiley, New York.

[1977] "The geometry of the set of controllable linear systems", Research reports of the Automatic Control Laboratory, Nagoya University, 24: 1-7.

P. BRUNOVSKY

[1970] "A classification of linear controllable systems", Kybernatika, 3: 173-187.

C. BYRNES and M. GAUGER

[1977] "Decidability criteria for the similarity problem, with applications to the moduli of linear dynamical systems", Advances in Mathematics, 25: 59-90.

J. CANDY

[1975] "Realization of invariant system descriptions from Markov sequences", Ph.D. dissertation, University of Florida, 124 pages.

J. CASTI

[1974] "Matrix Riccati equations, dimensionality reductions, and generalized X-Y functions", Utilitas Math., 6: 95-110.

J. FOGARTY

[1965] Invariant Theory, Benjamin, New York.

P. A. FUHRMANN and J. C. WILLEMS

[1979] "Factorization indices at infinity for rational matrix functions", Mathematisch Centrum, Amsterdam, 14 pages.

F. R. GANTMACHER

[1959] The Theory of Matrices, Volume 1, Chelsea, New York.

I. C. GOHBERG and M. G. KREIN

[1960] "Systems of integral equations on a half line with kernels depending on the difference of arguments", American Mathematical Society Translations, 14: 217-287.

M. HAZEWINKEL

[1977] "Moduli and canonical forms for linear dynamical systems III: the algebraic-geometric case", Proc. of the 1976 NASA-AMES Conference on Geometric Control Theory, Math. Sci. Press.

[1980] "A partial survey of the uses of algebraic geometry in systems and control theory", to appear in Symp. Math. INOAM (Severi Centennial Conference), Academic Press.

M. HAZEWINKEL and R. E. KALMAN

[1976] "On invariants, canonical forms, and moduli for linear, constant, finite dimensional, dynamical systems", Lecture Notes in Economics and Mathematical Systems Theory 131, pp. 48-60, Springer-Verlag, Berlin.

E. A. JONCKHEERE and L. M. SILVERMAN

[1978] "Spectral theory of linear-quadratic optimal control problem: discrete-time single-input case", IEEE Trans. on Circuits and Systems, CAS-25: 810-825.

T. KAILATH

[1973] "Some new algorithms for recursive estimation in constant linear systems", IEEE Trans. on Information Theory, IT-19: 750-760.

R. E. KALMAN

[1960] "Contributions to the theory of optimal control", Bol. Soc. Mat. Mex., 5: 102-119.

[1963] "On a new characterization of linear passive systems", Proceedings of the First Allerton Conference on Circuit and System Theory, pages 456-470.

[1964] "When is a linear system optimal?", Trans. ASME J. Basic Eng., Ser. D, 86: 51-60.

[1972] "Kronecker invariants and feedback", Proc. of NRL-MRC Conference on Ordinary Differential Equations (edited by L. Weiss), pages 459-471.

[1974] "Global structure of the class of linear dynamical systems of fixed dimension", Proc. 8th Princeton Conference on Information Sciences and Systems.

[1975] "Canonical forms for quadratic optimization problems", Research Proposal, Unpublished.

[1979] "On partial realizations, transfer functions, and canonical forms", Acta Polytechnica Scandinavica, Mathematics and Computer Science Series No. 31, pages 9-32.

R. E. KALMAN, Y. C. HO, and K. S. NARENDRA

[1963] "Controllability of linear dynamical systems", Contributions to Differential Equations, 1: 189-213.

B. P. MOLINARI

[1975] "Conditions for nonpositive solutions to linear matrix inequality", IEEE Transactions on Automatic Control, AC-20: 804-806.

D. MUMFORD

[1965] Geometric Invariant Theory, Springer-Verlag, Berlin.

R. W. NEWCOMB

[1966] Linear Multiport Synthesis, McGraw-Hill, New York.

V. M. POPOV

[1964] "Hyperstability and optimality of automatic systems with several control functions", Rev. Roum. Sci. Tech., Elektrotek. et Energ., 9: 629-690.

[1972] "Invariant description of linear time-invariant controllable systems", SIAM J. Control, 10: 252-264.

[1973] Hyperstability of Control Systems, Springer, Heidelberg.

C. PROCESI

[1976] "The invariant theory of $n \times n$ matrices", Advances in Mathematics, 19: 306-381.

M. ROSENBLUM

[1965] "A concrete spectral theory for self adjoint Toeplitz operators", American J. Mathematics, 87: 709-718.

H. H. ROSENBROCK

[1970] State Space and Multivariable Theory, Nelson, London.

K. S. SIBIRSKII

[1967] "On unitary and orthogonal invariants", Soviet Math. Dokl., 8: 36-40.

L. M. SILVERMAN

[1976] "Discrete Riccati equations: alternative algorithms, asymptotic properties, and system theory interpretations", in Control and Dynamical Systems, Vol. 12 (edited by C. T. Leondes), Academic Press, New York.

A. TANNENBAUM

[1980] Invariance and System Theory: Algebraic and Geometric Aspects, to appear in the Springer Lecture Notes in Mathematics, 81 pages.

S. H. WANG and E. J. DAVISON

[1976] "Canonical forms for linear multivariable systems," SIAM J. Control and Optimization, 14: 236-250.

J. C. WILLEMS

[1971] "Least squares stationary optimal control and the algebraic Riccati equation", IEEE Transactions on Automatic Control, AC-16: 621-634.

[1981] Linear Optimal Control and Filtering, forthcoming.

W. M. WONHAM

[1974] Linear Multivariable Control: A Geometric Approach, Lecture Notes in Economics and Mathematical Systems No. 101, Springer, Berlin.

V. A. YAKUBOVIC

[1962] "The solution of certain matrix inequalities in automatic control theory", Dokl. Akad. Nauk, SSSR, 143: 1304-1307.

K. YOSIDA

[1965] Functional Analysis, Springer, Berlin.

BIOGRAPHICAL SKETCH

Pramod Prabhakar KHARGONEKAR was born on August 24, 1956, in Indore, INDIA, to Prabhakar KHARGONEKAR and Leela KHARGONEKAR. He obtained his Bachelor of Technology degree in electrical engineering from the Indian Institute of Technology, Bombay in 1977. He obtained his Master of Science degree in mathematics in 1980 from the University of Florida.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

R E Kalman

Dr. R. E. Kalman, Chairman
Graduate Research Professor of
Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

T. E. Bullock

Dr. T. E. Bullock, Co-Chairman
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert L. Long

Dr. R. L. Long
Assistant Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

V. M. Popov

Dr. V. M. Popov
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Charles V. Shaffer

Dr. C. V. Shaffer
Professor of Electrical Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1981.

Wayne H. Chen

Dean, College of Engineering

Dean, Graduate School

UNIVERSITY OF FLORIDA
3 1262 08394 207 7